

94473

From: Rao, Manjunath N.
Sent: Tuesday, May 20, 2003 11:54 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/869,155

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10D 01
Phone: 306-5681

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

Date: 5-20-03

Please search the following as soon as possible for application with serial number **09/869,155**

1. SEQ ID NO:13, 14, 15, 16, 17, 18, 19 against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Mail Box in 10D01
Crystal Mall 1, USPTO.

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 5/20
Date Completed: 5/20
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 : Search time 8.40659 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-14

Perfect score: 96

Sequence: 1 KRLGSRRLPHFTGCGL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCPTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.5	46.4	218	9	US-09-738-626-4068
2	43	44.8	163	9	US-10-211-962-76
3	42	43.8	163	9	US-10-211-962-71
4	41	42.7	163	9	US-10-211-962-72
5	41	42.7	163	9	US-10-211-962-75
6	41	42.7	729	9	US-10-051-902-29
7	41	42.7	729	9	US-10-051-909-29
8	40	41.7	163	9	US-10-211-962-79
9	40	41.7	324	10	US-09-816-095-4
10	40	41.7	341	9	US-10-043-487-290
11	40	41.7	372	9	US-09-860-670-106
12	39	40.6	59	9	US-09-796-692-1038
13	39	40.6	59	9	US-10-040-862-1038
14	39	40.6	64	9	US-10-150-111-155
15	39	40.6	572	9	US-10-272-419-8
16	38	39.6	25	9	US-10-023-282-454
17	38	39.6	85	9	US-10-114-893-194
18	38	39.6	123	10	US-09-867-550-1202
19	38	39.6	197	10	US-09-764-853-774

20	38	39.6	363	9	US-09-738-626-5456	Sequence 5456, App
21	38	39.6	499	10	US-09-976-165-40	Sequence 40, App1
22	38	39.6	620	10	US-09-925-301-1193	Sequence 1193, App
23	38	39.6	737	9	US-10-051-902-8	Sequence 8, App11
24	38	39.6	737	9	US-10-051-909-8	Sequence 8, App11
25	38	39.6	782	9	US-10-097-340-312	Sequence 312, App
26	38	39.6	855	9	US-10-099-700A-2	Sequence 2, App11
27	38	39.6	1189	9	US-09-738-626-4140	Sequence 4140, App
28	38	39.6	5215	9	US-09-860-846-2	Sequence 2, App11
29	38	39.6	5215	9	US-09-888-384B-2	Sequence 2, App11
30	38	39.6	5215	9	US-09-836-821-2	Sequence 2, App11
31	38	39.6	5215	10	US-09-861-289-2	Sequence 2, App11
32	37.5	39.1	124	9	US-09-764-891-4424	Sequence 4424, App
33	37.5	39.1	402	10	US-09-815-242-5748	Sequence 5748, App
34	37.5	39.1	408	10	US-09-815-242-12296	Sequence 12296, App
35	37.5	39.1	669	9	US-10-024-632-9	Sequence 9, App11
36	37	38.5	67	9	US-09-796-692-752	Sequence 752, App
37	37	38.5	67	9	US-10-040-862-752	Sequence 752, App
38	37	38.5	80	9	US-09-796-692-889	Sequence 889, App
39	37	38.5	80	9	US-10-040-862-889	Sequence 889, App
40	37	38.5	120	9	US-09-796-692-807	Sequence 807, App
41	37	38.5	120	9	US-10-040-862-807	Sequence 807, App
42	37	38.5	295	10	US-09-815-242-11239	Sequence 11239, App
43	37	38.5	567	10	US-09-815-242-13511	Sequence 13511, App
44	37	38.5	740	9	US-10-051-909-37	Sequence 37, App1
45	37	38.5	800	9	US-10-051-909-32	Sequence 32, App1

ALIGNMENTS

RESULT 1
US-09-738-626-4068
; Sequence 4068, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KETIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4068
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4068

Query Match 46.4%; Score 44.5; DB 9; Length 218;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
OY 2 KRLGSRRLPHFTGCGL 17
DB 76 RQGSIRAPHTG-GGI 90

```
RESULT 2
US-10-211-962-76
; Sequence 76, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-76
```

```
Query Match          44.8%; Score 43; DB 9; Length 163;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 KRIGFSRLPHFTG 12
    | | | | | | | | | |
Db 84 KMGFGRLHFHS 95
```

```
RESULT 3
US-10-211-962-71
; Sequence 71, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-71
```

```
Query Match          43.8%; Score 42; DB 9; Length 163;
Best Local Similarity 61.5%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 KRIGFSRLPHFTG 13
    | | | | | | | | | |
Db 84 KMGFGRLHFHS 96
```

```
RESULT 4
US-10-211-962-72
; Sequence 72, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
```

```
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-72
```

```
Query Match          42.7%; Score 41; DB 9; Length 163;
Best Local Similarity 61.5%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 KRIGFSRLPHFTG 13
    | | | | | | | | | |
Db 84 KMGFGRLHFHS 96
```

```
RESULT 5
US-10-211-962-75
; Sequence 75, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-75
```

```
Query Match          42.7%; Score 41; DB 9; Length 163;
Best Local Similarity 61.5%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
OY 1 KRIGFSRLPHFTG 13
    | | | | | | | | | |
Db 84 KMGFGRLHFHS 96
```

```
RESULT 6
US-10-051-902-29
; Sequence 29, Application US/10051902
; Patent No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
```



```

; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-902-29

Query Match
Best Local Similarity 42.7%; Score 41; DB 9; Length 729;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LPHFTGCGGL 17
DB 136 LPHFTGCGGM 145

RESULT 7
US-10-051-909-29
; Sequence 29, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: Bb1163 US.CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-909-29

Query Match
Best Local Similarity 42.7%; Score 41; DB 9; Length 729;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LPHFTGCGGL 17
DB 136 LPHFTGCGGM 145

RESULT 8
US-10-211-962-79
; Sequence 79, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
```

```

US-10-211-962-79

Query Match
Best Local Similarity 41.7%; Score 40; DB 9; Length 163;
Best Local Similarity 61.5%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KRGFSLPHFTG 13
DB 84 KMGFLRLHFHSG 96

RESULT 9
US-09-816-095-4
; Sequence 4, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-816-095-4

Query Match
Best Local Similarity 41.7%; Score 40; DB 10; Length 324;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PHFTGCGG 16
DB 42 PHFTGCGG 49

RESULT 10
US-10-043-487-290
; Sequence 290, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypep
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 290
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-290

Query Match
Best Local Similarity 41.7%; Score 40; DB 9; Length 341;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GFSRLPHFTGCGGL 17
DB 291 GFSRLPHFTGCGGM 304

RESULT 11
US-09-860-670-106
```

```
Sequence 106, Application US/09860670
Patent No. US20020165137A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA127P1
CURRENT APPLICATION NUMBER: US/09/860,670
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 289
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 106
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-09-860-670-106

Query Match          41.7%; Score 40; DB 9; Length 372;
Best Local Similarity 47.8%; Pred. No. 1.4e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 6; Gaps 2;

QY 1 KRIGFSR---LPHF---TGGGCL 17
   11:11:1 1 1 1 1 1 1
Db 56 KRIGFMSYVLAFTFAAGCAGL 78

RESULT 12
US-09-796-692-1038
Sequence 1038, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Aigate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1038
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-692-1038

Query Match          40.6%; Score 39; DB 9; Length 59;
Best Local Similarity 58.3%; Pred. No. 34;
```

```
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LGFSRLPHFTGC 14
   11:11111 1
Db 18 LGSAYVLPHPFSC 29

RESULT 13
US-10-040-862-1038
Sequence 1038, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Aigate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1038
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-10-040-862-1038

Query Match          40.6%; Score 39; DB 9; Length 59;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LGFSRLPHFTGC 14
   11:11111 1
Db 18 LGSAYVLPHPFSC 29

RESULT 14
US-10-150-111-155
Sequence 155, Application US/10150111
Publication No. US20030078386A1
GENERAL INFORMATION:
APPLICANT: Rubin et al.
TITLE OF INVENTION: Secreted Protein HPEAD48
FILE REFERENCE: P2018P1D1
```

;; CURRENT APPLICATION NUMBER: US/10/150,111
;; CURRENT FILING DATE: 2002-05-20
;; PRIOR APPLICATION NUMBER: 09/288,143
;; PRIOR FILING DATE: 1999-04-08
;; PRIOR APPLICATION NUMBER: PCT/US98/21142
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/061,463
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: 60/061,529
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: 60/071,498
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: 60/061,527
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: 60/061,536
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: 60/061,532
;; PRIOR FILING DATE: 1997-10-09
;; NUMBER OF SEQ ID NOS: 219
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 155
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-150-111-155

Query Match 40.6%; Score 39; DB 9; Length 64;
Best Local Similarity 47.6%; Pred. No. 37;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

OY 1 KRIGSRPLPHT---GCGGL 17
I :||| ||:|
Db 16 KWIGFSIFPHWTWIDLEIGL 36

RESULT 15
US-10-272-419-8
;; Sequence 8, Application US/10272419
;; Publication No. US20030087403A1
;; GENERAL INFORMATION:
;; APPLICANT: CHEN, QIONG
;; APPLICANT: THOMAS, STUART
;; APPLICANT: NAGARAJAN, VASANTHA
;; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
;; FILE REFERENCE: C11341-A
;; CURRENT APPLICATION NUMBER: US/10/272,419
;; CURRENT FILING DATE: 2002-10-16
;; PRIOR APPLICATION NUMBER: 09/252,553
;; PRIOR FILING DATE: 1999-02-19
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO: 8
;; LENGTH: 572
;; TYPE: PRT
;; ORGANISM: Acinetobacter sp.
US-10-272-419-8

Query Match 40.6%; Score 39; DB 9; Length 572;
Best Local Similarity 61.5%; Pred. No. 31e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 GFSRPLPHTGCG 16
I :||| ||:|
Db 316 GFNSLPKFTTPG 328

Search completed: May 20, 2003, 18:00:00
Job time : 8.40659 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 : Search time 38.0769 seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-13
Perfect score: 185
Sequence: 1 GAPVRAVEAVAPGVCYDKTKLGNLGVAVPVV 35

Scoring table:
BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	100.0	35	21	AAE93756
2	185	100.0	35	22	AAE14687
3	185	100.0	35	22	AAU07392
4	185	100.0	196	23	AAU75817
5	185	100.0	370	23	AAU75815
6	185	100.0	370	23	AAU75816
7	180	97.3	381	23	AAU75812
8	175	94.6	381	23	AAU75813
9	170	91.9	381	23	AAU75814
10	162	87.6	185	23	AAU75824

11	145	78.4	57	21	AAV93763
12	141	76.2	185	23	AAU75826
13	139	75.1	185	23	AAU75825
14	111	60.0	102	23	AAU75820
15	105	56.8	102	23	AAU75821
16	99	53.5	277	23	AAU75818
17	97	52.4	170	23	AAU75827
18	94	50.8	20	23	AAU75809
19	87	47.0	287	23	AAU75819
20	85	45.9	25	23	AAU75810
21	80.5	43.5	258	21	AAU44139
22	80.5	43.5	433	21	AAU44138
23	80.5	43.5	439	21	AAU44137
24	79.5	43.0	258	21	AAU42492
25	79.5	43.0	433	21	AAU42491
26	79.5	43.0	438	21	AAU42490
27	66	35.7	405	21	AAU44494
28	66	35.7	425	21	AAU44493
29	56.5	30.5	525	22	AAU34656
30	56.5	30.5	525	22	AAU36148
31	56.5	30.5	525	22	AAU38357
32	54	28.2	24	23	AAU75811
33	54	28.2	546	19	AAU39788
34	53.5	28.9	311	23	ABB54685
35	53	28.6	152	20	AAU74030
36	52.5	28.4	525	22	AAU36359
37	52	28.1	142	22	AAU92646
38	52	28.1	212	15	AAU45357
39	52	28.1	212	21	AAU16445
40	52	27.8	486	22	AAU42916
41	51.5	27.8	516	20	AAU15908
42	51.5	27.8	520	22	AAU37690
43	51	27.6	707	22	ABG24674
44	50	27.0	444	23	ABP26884
45	50	27.0	534	17	AAU91309

ALIGNMENTS

RESULT 1	AAV93756	standard; peptide: 35 AA.
ID	AAV93756;	
AC	AAV93756;	
XX		
XX		
DT	03-OCT-2000 (first entry)	
DE	Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.	
XX		
KW	Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product; dough; dough preparation.	
KW		
XX		
OS	Triticum sp.	
XX		
PN	WO200039289-A2.	
XX		
PD	06-JUL-2000.	
XX		
PF	17-DEC-1999; 99WO-IB02071.	
XX		
PR	23-DEC-1998; 98GB-0028599.	
PR	06-APR-1999; 99GB-0007805.	
PR	15-APR-1999; 99GB-0008645.	
PA	(DANI-) DANISCO AS.	
XX		
PI	Sibbesen O, Sorensen JF;	
XX		
DR	WPI; 2000-465744/40.	
XX		
PT	Mutant xylanase protein identified using xylanase inhibitor useful for preparing non-sticky dough for bakery products	

XX Claim 24: Page 112; 112pp; English.
PS
XX The present sequence is derived from an endo-beta-1,4-xyylanase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xyylanase protein. The xyylanase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xyylanase or
CC mutant xyylanase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xyylanase. The xyylanase inhibitor is
CC useful for screening high degree resistance xyylanses for dough
CC preparation. The xyylanase is also useful for preparing a non-sticky
CC dough. A combination of xyylanase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
SQ Sequence 35 AA:

Query Match 100.0%; Score 185; DB 21; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.3e-21;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GAPVARAVEAVAPFGVCYDTKTLGNLGGYAVPNV 35
Db 1 GAPVARAVEAVAPFGVCYDTKTLGNLGGYAVPNV 35
|||||

RESULT 2
AAEI4687
ID AAEI4687 standard: peptide; 35 AA.
XX
XX AAEI4687;
AC
XX
XX 21-AUG-2002 (first entry)
DT
XX
XX Wheat flour xyylanase inhibitor A chain N-terminal fragment.
DE
XX Refrigerated dough; syrruping; arabinoxylan; bakery product; bread;
XX pizza base; cake; biscuit; wheat; flour; xyylanase inhibitor.
KW
XX Triticum aestivum.
OS
XX WO200152657-A1.
PM
XX 26-JUL-2001.
PC
XX 17-JAN-2001; 2001WO-1B00168.
XX
XX 18-JAN-2000; 2000GB-0001136.
PR
XX
XX (DANI-) DANISCO AS.
PA
XX
XX Poulsen CH, Sorensen JF;
PI
XX
XX WPI: 2001-457446/49.
DR
XX
XX Production of refrigerated dough with reduced syrruping, useful in
PT production of bakery products such as bread, comprises admixing cereal
PT flour, water and protein that prevents enzymatic degradation of
PT arabinoxylan in the cereal flour -
XX
XX
XX Disclosure: Page 23; 26pp; English.
PS
XX
XX The invention relates to a process for producing refrigerated dough
CC with reduced 'syruping' (precipitation of liquid on the dough surface
CC because of a reduction in water holding capacity caused by the breakdown
CC of arabinoxylan over time). The process comprises admixing cereal flour
CC and water with a protein that reduces/prevents enzymatic degradation of
CC arabinoxylan in the cereal flour. The preferred protein is a xyylanase
CC inhibitor. The method is useful to produce refrigerated dough in which
CC syruping is reduced or eliminated. Refrigerated dough is typically
CC stored for long periods to enable fresh baked products (e.g. bread,
CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
CC requirements using the method by the use of specific proteins/protein
CC combinations. The present sequence is wheat flour
CC endo-beta-1,4-xyylanase inhibitor A chain N-terminal fragment.
XX
SQ Sequence 35 AA:

Query Match 100.0%; Score 185; DB 22; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.3e-21;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAPVARAVEAVAPFGVCYDTKTLGNLGGYAVPNV 35
Db 1 GAPVARAVEAVAPFGVCYDTKTLGNLGGYAVPNV 35
|||||

RESULT 3
AAU07392
ID AAU07392 standard: protein; 35 AA.
XX
XX AAU07392;
AC
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Bacillus subtilis xyylanase inhibitor #1.
DE
XX
XX xyylanase; plant cell wall; baking; cereal; starch production; wood;
KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
XX
XX Bacillus subtilis.
OS
XX WO200166711-A1.
PM
XX
XX 13-SEP-2001.
PD
XX
XX 08-MAR-2001; 2001WO-1B00426.
PF
XX
XX 08-MAR-2000; 2000GB-0005585.
PR
XX 27-JUN-2000; 2000GB-0015751.
XX
XX (DANI-) DANISCO AS.
PA
XX
XX Sidsesen O, Sorensen JF;
PI
XX
XX WPI: 2001-596834/67.
DR
XX
XX Novel variant xyylanase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xyylanase inhibitor -
XX
XX
XX Disclosure: Page 63; 70pp; English.
PS
XX
XX The invention relates to a variant xyylanase polypeptide (I) or its
CC fragment having xyylanase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xyylanase inhibitor as compared with the parent xyylanase enzyme. (I) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I)
CC is useful for altering the viscosity derived from the presence of
CC hemicellulose or arabinoxylan in a solution or system comprising plant
CC cell wall material. (I) is useful for preparing a foodstuff such as
CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
CC The present sequence represents the amino acid sequence of Bacillus
CC subtilis xyylanase inhibitor #1 as described in the method of the
CC invention.
XX
SQ Sequence 35 AA:

Query Match 100.0%; Score 185; DB 22; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.3e-21;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAPVARAEAVAPFGVCYDRTKLTGNNLGGYAVPNV 35
 |||
 DB 1 GAPVARAEAVAPFGVCYDRTKLTGNNLGGYAVPNV 35

RESULT 4
 AAU75817
 ID AAU75817 standard; Protein: 196 AA.
 AC AAU75817;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Wheat L endoxylanase inhibitor, TDXI I, partial sequence TDXI-I.01.
 XX
 KW Wheat; TDXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 OS Triticum durum cultivar Mexicali.
 XX
 XX MO200198474-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEUVEN RES & DEV.
 PA
 XX Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX
 DR WPI; 2002-114579/15.
 DR N-PSDB; ABK13674.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruiping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TDXI-I (T. durum L endoxylanase
 CC inhibitor).
 CC
 XX
 SQ Sequence 196 AA;

Query Match 100.0%; Score 185; DB 23; Length 196;
 Best Local Similarity 100.0%; Pred. No. 6; Le-20;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAPVARAEAVAPFGVCYDRTKLTGNNLGGYAVPNV 35
 |||
 DB 81 GAPVARAEAVAPFGVCYDRTKLTGNNLGGYAVPNV 115

RESULT 5
 AAU75815
 ID AAU75815 standard; Protein: 370 AA.
 AC AAU75815;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.
 XX
 KW Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 OS Triticum aestivum cultivar Soissons.
 XX
 XX MO200198474-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEUVEN RES & DEV.
 PA
 XX Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX
 DR WPI; 2002-114579/15.
 DR N-PSDB; ABK13672.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitor
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC arabinoxylan/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).

CC Sequence 370 AA:

Query Match 100.0%; Score 185; DB 23; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.3e-19;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAPVAVAEVAPFGVCYDTKTLGNLGGYAVPNV 35
 ID 255 GAPVAVAEVAPFGVCYDTKTLGNLGGYAVPNV 289

RSULT 6
 ID AAU75816 standard; Protein: 370 AA.

AC AAU75816;

DT 23-APR-2002 (first entry)

DE Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.02.

KW Wheat, TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal products; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Triticum aestivum cultivar Estica.

PN W0200198474-A1.

PD 27-DEC-2001.

PE 21-JUN-2001; 2001WO-BE00106.

PR 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.

PR 16-MAR-2001; 2001GB-0006564.

PR 21-MAY-2001; 2001GB-0012328.

XX (LEUV-) LEUVEN RES & DEV.

XX Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX WPI: 2002-114579/15.
 DR N-PSDB; AKR13673.
 XX
 PR Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PI beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes

Claim 127; Page 57; 127pp; English.

The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitor
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).

Sequence 370 AA:

Query Match 100.0%; Score 185; DB 23; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.3e-19;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAPVAVAEVAPFGVCYDTKTLGNLGGYAVPNV 35
 ID 255 GAPVAVAEVAPFGVCYDTKTLGNLGGYAVPNV 289

RSULT 7
 ID AAU75812 standard; Protein: 381 AA.

AC AAU75812;

DT 23-APR-2002 (first entry)

DE Wheat L endoxylanase inhibitor, TAXI I, variant #1.

KW Wheat, TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal products; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX Triticum aestivum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95 /Label= unknown
 FT Misc-difference 98 /Label= unknown
 FT Misc-difference 101 /Label= unknown
 FT Misc-difference 110 /Label= unknown
 FT Misc-difference 333 /Label= unknown
 FT Misc-difference /Label= unknown
 XX WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 23-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI: 2002-114579/15.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 50-51; 127pp; English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, maltose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-I ('T. aestivum L endoxylanase

CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 SQ Sequence 381 AA:
 Query Match 97.3%; Score 180; DB 23; Length 381;
 Best Local Similarity 97.1%; Pred. No. 8e-19;
 Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAPVARAVEAVAPFCGYCTKTLGNLGGYAVPNV 35
 DB 266 GAPVARAVEVAPFVGVCYDPTKTLGNLGGYAVPNV 300
 RESULT 8
 AAU75813
 ID AAU75813 standard; Protein; 381 AA.
 AC
 XX AAU75813;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Wheat L endoxylanase inhibitor, TAXI I, variant #2.
 XX
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KW
 XX Triticum aestivum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95 /Label= unknown
 FT Misc-difference 98 /Label= unknown
 FT Misc-difference 101 /Label= unknown
 FT Misc-difference 110 /Label= unknown
 FT Misc-difference 333 /Label= unknown
 FT Misc-difference /Label= unknown
 XX WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 23-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI: 2002-114579/15.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 51; 127pp; English.
 CC The invention relates to separating and/or isolating inhibitors of

cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabinofuranosidase and/or other cellulose, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syruing in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylanases such as by the formation of an endoxylanase/inhibitor complex, improving the malting of cereals such as barley, sorghum and wheat and/or the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylitol and/or xylitol, the wheat gluten starch separation and production, improving maize processing, plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence is a variant of wheat TAXI-1 (T. aestivum L endoxylanase inhibitor).

Note: Variant amino acids are highlighted in the specification but no wild-type sequence is shown for comparison.

Misc-difference 145
 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 Misc-difference 183
 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 Misc-difference 232
 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 Misc-difference 275
 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 Misc-difference 282
 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 Misc-difference 333
 /label= Unknown
 WO200198474-A1.
 27-DEC-2001.
 21-JUN-2001; 2001WO-BE00106.
 22-JUN-2000; 2000GB-0015296.
 25-JAN-2001; 2001GB-0002018.
 26-JAN-2001; 2001GB-0002194.
 16-MAR-2001; 2001GB-0006564.
 21-MAY-2001; 2001GB-0012328.
 (LEUV-) LEUVEN RES. & DEV.
 Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 Van Campenhout S;
 WPI; 2002-114579/15.
 Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 beta-glucanolytic enzymes comprises using endoxylanases during
 screening for inhibition activity or affinity chromatography with

	Query Match	94.6%	Score 175;	DB 23		Length 381;
	Best Local Similarity	97.1%;	Pred. No.	4.6e-18;		
	Matches 34;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
DQ	1 GAPVARAAVEAAPFCGYDTRKTLGNLGGYAVPNV 35					
DQ	266 GAPVARAAVEAAPFCGLYDTRKTLGNLGGYAVPNV 300					
RESULT 9						
ID	AAU75814					
XX	AAU75814 standard; Protein; 381 AA.					
AC	AAU75814;					
XX						
DT	23-APR-2002 (first entry)					
DE	Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.					
XX						
KM	Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;					
KM	xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;					
KM	amylolytic enzyme; enzyme; dough syruping; cereal product; beer; plant;					
KM	straight dough; sponge dough; Chorleywood bread; biscuits; pasta;					
KM	noodle; animal feed; starch separation; maize processing; malting;					
KM	plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.					
XX						
OS	Triticum aestivum.					
XX						
FH	Key Location/Qualifiers					
FT	Misc-difference 95 /label= Unknown					
FT	Misc-difference 98 /label= Unknown					
FT	Misc-difference 101 /label= Unknown					
FT	Misc-difference 110 /label= Unknown					
FT	/label= Unknown					

XX Sequence 381 AA;
SQ Query Match 91.9%; Score 170; DB 23; Length 381;
Best Local Similarity 94.3%; Pred. No. 2,7e-17;
Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GAPVAAVEAVAPFGVCYDTKLTGNNLGGYAVPNV 35
Db 266 GAPVAAVEAVAPFGVCYDTKLTGNNLGGYAVPNV 300
RESULT 10
AAU75824
ID AAU75824 standard; Protein: 185 AA.
XX AAU75824;
AC
XX 23-APR-2002 (first entry)
XX
DE Barley L endoxylanase inhibitor, HVXI I, variant #1.
XX
KW Barley; HVXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
OS Hordeum vulgare.
FH Key Location/Qualifiers
FT Misc-difference 72 /label= Unknown
FT Misc-difference 185 /label= Unknown
FT
XX WO200198474-A1.
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
XX 25-JAN-2001; 2001GB-0002018.
XX 26-JAN-2001; 2001GB-0002194.
XX 16-MAR-2001; 2001GB-0006564.
XX 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES & DEV.
XX
XX Delcours J, Debysse W, Gebuere K, Goesaert H, Flerens K, Robben J;
PI Van Campenhout S;
XX
XX WPI: 2002-114579/15.
XX
XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT beta-glucanolytic enzymes comprises using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
XX Claim 127; Page 63; 127pp; English.
XX
XX The invention relates to separating and/or isolating inhibitors of
XX cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
XX screening the inhibition activity by using two or more enzymes during the
XX separation and/or isolation steps that allow to distinguish inhibitors of
XX different specificity or by using an affinity chromatographic step with
XX immobilised enzymes and/or antibodies against inhibitors. Also
XX included are an isolated nucleic acid molecule encoding an inhibitor
XX which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
XX alpha-L-arabino-furanosidase and/or other cellulose, xylan,
XX arabinoxylan or beta-glucan degrading enzymes, a vector comprising the

CC nucleic acid, an expression system transformed with the nucleic
CC acid, a host organism transformed with the nucleic acid, the inhibitory
CC proteins encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syruping in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specificity and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC endoxylanase/inhibitor complex, improving the malting of cereals such as
CC barley, sorghum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch separation and production, improving maize processing,
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence is a variant of barley HVXI-I (H. vulgare L endoxylanase
CC inhibitor).
CC Note: Variant amino acids are highlighted in the specification but
CC no wild-type sequence is shown for comparison.
XX
SQ Sequence 185 AA;
Query Match 87.6%; Score 162; DB 23; Length 185;
Best Local Similarity 82.9%; Pred. No. 1.8e-16;
Matches 29; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 GAPVAAVEAVAPFGVCYDTKLTGNNLGGYAVPNV 35
Db 126 GAPVAAVEAVAPFGVCYDTKLTGNNLGGYAVPNV 160
RESULT 11
AA93763
ID AA93763 standard; peptide; 57 AA.
XX
XX AA93763;
AC
XX 03-OCT-2000 (first entry)
XX
XX Amino acid sequence derived from an endo-beta-1,4-xylanase inhibitor.
XX
XX Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
XX dough; dough preparation.
XX
XX Triticum sp.
XX
XX
XX Key Location/Qualifiers
FH Misc-difference 17 /note= "unspecified amino acid"
FT Misc-difference 43 /note= "unspecified amino acid"
FT Misc-difference 49 /note= "unspecified amino acid"
FT
XX WO200039289-A2.
XX
XX 06-JUL-2000.
XX
XX 17-DEC-1999; 99WO-IB02071.
XX
XX 23-DEC-1998; 98GB-0028599.
XX 06-APR-1999; 99GB-0007805.
XX 15-APR-1999; 99GB-0008645.
XX
XX (DANI-) DANISCO AS.
XX
XX Slibesen O, Sorensen JF;
PI
XX

DR	WPI: 2000-465744/40.
XX	
PT	Mutant xylanase protein identified using xylanase inhibitor useful for
PT	preparing non-sticky dough for bakery products
XX	
PS	Disclosure: Page 106; 112pp: English.
XX	
CC	The present sequence is derived from an endo-beta-1,4-xylanase
CC	inhibitor. The protein is obtained from wheat flour. The specification
CC	also describes a mutant xylanase protein. The xylanase is useful for
CC	preparing a foodstuff, preferably a bakery product or a substance
CC	(e.g. a dough) for making the bakery product. Wild type xylanase or
CC	mutant xylanase is useful for preparing a dough that is less sticky
CC	than a dough comprising a fungal xylanase. The xylanase inhibitor is
CC	useful for screening high degree resistance xylanases for dough
CC	preparation. The xylanase is also useful for preparing a non-sticky
CC	dough. A combination of xylanase and the inhibitor is useful for
CC	calibrating and/or determining the quantity of inhibitor in a wheat
CC	flour sample.
XX	
EQ	Sequence 57 AA:
	Query Match 78.4%; Score 145; DB 21; Length 57;
	Best Local Similarity 87.9%; Pred. No. 1.6e-14;
	Matches 29; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY	2 APVARAEAVAFPGYCYDTKLTGNMLGAYAVN 34
	1 : : : :
D3	2 AVVARAVKVDVAFGVXYDTKLTGNMLGAYAVN 34
RESULT 12	
AAU075826	
ID	AAU75826 standard; Protein: 185 AA.
XX	
AC	AAU75826;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Barley L endoxylanase inhibitor, HVX1 I, microheterogenic variant.
XX	
KW	Barley: HVX1-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW	xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW	immobilised enzyme; enzyme; dough; syruup; cereal product; beer; plant;
KW	straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW	noodle; animal feed; starch separation; maize processing; malting;
KW	plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX	
OS	Hordeum vulgare.
OS	Synthetic.
XX	
EH	Key Location/Qualifiers
FT	Misc-difference 72
FT	/label= Unknown
FT	Misc-difference 128
FT	/label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
FT	Misc-difference 132
FT	/label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
FT	Misc-difference 134
FT	/label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
FT	Misc-difference 148
FT	/label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
FT	Misc-difference 185
FT	/label= Unknown
XX	
PN	MO200198474-A1.
XX	
PD	27-DEC-2001.
XX	
PF	21-JUN-2001; 2001WO-BE00106.
XX	
RR	22-JUN-2000; 2000GB-0015296.
PR	25-JAN-2001; 2001GB-0002018.

PR		26-JAN-2001; 2001GB-0002194.
PR		16-MAR-2001; 2001GB-0006564.
PR		21-MAY-2001; 2001GB-0012328.
PA	(LEUV-) LEUVEN RES & DEV.	
PI	Delfour J, Debysse W, Gebruers K, Goesaert H, Flierens K, Robben J;	
PI	Van Campenhout S;	
DR	WPI: 2002-114579/15.	
PT	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or	
PT	beta-glucanolytic enzymes comprises using endoxylanases during	
PT	screening for inhibition activity or affinity chromatography with	
PT	immobilised enzymes	
PS	Claim 127; Page 63; 127pp; English.	
XX		
CC	The invention relates to separating and/or isolating inhibitors of	
CC	cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising	
CC	screening the inhibition activity by using two or more enzymes during the	
CC	separation and/or isolation steps that allow to distinguish inhibitors of	
CC	different specificity or by using an affinity chromatographic step with	
CC	immobilised enzymes and/or antibodies against inhibitors. Also	
CC	included are an isolated nucleic acid molecule encoding an inhibitor	
CC	which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,	
CC	alpha-L-arabino-furanosidase and/or other cellulose, xylan,	
CC	arabinoxylan or beta-glucan degrading enzymes, a vector comprising the	
CC	nucleic acid, an expression system transformed with the nucleic	
CC	acid, a host organism transformed with the nucleic acid, the inhibitory	
CC	proteins encoded by the nucleic acids and modulators of the proteins.	
CC	A recombinant protein, glycoprotein or polypeptide or microorganisms,	
CC	plant or plant materials transformed with the nucleic acid are useful	
CC	for the formation of an endoxylanase-inhibitor complex, screening	
CC	endoxylanases that are totally, less or not inhibited by the inhibitors,	
CC	reducing syruiping in refrigerated dough compositions, affecting the	
CC	relative affinity and/or relative hydrolysis specificity and/or relative	
CC	hydrolysis rate versus water-extractable and/or water-unextractable	
CC	arabinoxylans of endoxylanases such as by the formation of cereal	
CC	endoxylanase/inhibitor complex, improving the malting of cereals such as	
CC	barley, sorghum and wheat and/or the production of beer, improving the	
CC	production and/or quality of baked or extruded cereal products such as	
CC	straight dough, sponge dough, Chorleywood bread, breakfast cereals,	
CC	biscuits, pasta and noodles, animal feed stuff, improving the production	
CC	of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten	
CC	starch separation and production, improving maize processing,	
CC	plant disease resistance and nutraceutical and/or pharmaceutical	
CC	applications, improving paper and pulp technologies. The present	
CC	sequence is a microheterogenic variant of barley HVXI-I (H. vulgare L	
CC	endoxylanase inhibitor).	
CC	Note: Variant amino acids are highlighted in the specification but	
CC	no wild-type sequence is shown for comparison.	
XX		
SQ	Sequence 185 AA:	
Query Match	76.2%; Score 141; DB 23; Length 185;	
Best Local Similarity	74.3%; Pred. NO.2,8e-13;	
Matches	26; Conservative 2; Mismatches 7; Indels 0; Gaps 0;	
OY	1 GAPVARAVEAVAFEGCYDTKTLGNNGGYAVPNV 35	
DB	126 GAXAARXYVPVAPFGLCYDAKTKXGNNLGGYSYPNV 160	
RESULT 13		
AAU75825		
ID	AAU75825 standard; Protein: 185 AA.	
AC	AAU75825;	
XX		
DT	23-APR-2002 (first entry)	
XX		
Barley L endoxylanase inhibitor, HVXI I, variant #2.		

XX Barley; HMXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX Hordeum vulgare.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /Label- Unknown
 FT Misc-difference 185 /Label- Unknown
 FT
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 PI
 DR WPI: 2002-114579/15.
 XX
 PT
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127: Page 63; 127pp: English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of Barley HMXI-I (H. vulgare L endoxylanase

CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 SQ Sequence 185 AA:
 Query Match 75.1%; Score 139; DB 23; Length 185;
 Best Local Similarity 74.3%; Pred. NO. 5,7e-13;
 Matches 26; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 GAPVAAVEAVAPFGVCYDTKLGNNLGGYAVPNV 35
 Db 126 GALAAGVNPVAPFGICYDAKTGNLGGYSVPNV 160
 RESULT 14
 AAU75820
 ID AAU75820 standard; Protein; 102 AA.
 XX
 AC AAU75820;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Rye L endoxylanase inhibitor, partial sequence SCXI-01.
 XX
 KW Rye; SCXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Secale cereale cultivar Halo.
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 PI
 DR WPI: 2002-114579/15.
 DR N-PSDB: ABK13677.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127: Page 59; 127pp: English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory

CC protein encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syripping in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specifically and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC endoxylanase/inhibitor complex, improving the malting of cereals such as
CC barley, sorghum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
CC starch separation and production, improving maize processing,
CC plant disease resistance and production, improving pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence represents a partial rye *SxR1-1* (S. cereale L endoxylanase
CC inhibitor).

Sequence 102 AA;

Query Match	60.0%;	Score 111;	DB 23;	Length 102;
Best Local Similarity	91.38;	Pred. No. 5e-09;		
Matches 21; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0

```
Qy      1 GAPVARAVEAVAPFGVCYDTKTL 23
          |||||: |||||
Db      80 GAPVARAVQPVAPFGVCYDTKTL 102
```

RESULT 15
AAU75821

ID AU75821 standard; Protein; 102 AA

AC AAU75821;

DT 23-APR-2002 (first entry)

DE Rye L endoxylanase inhibitor, partial sequence SCXI-02.

KW Rye, SCAR1: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW amylohydric enzyme; enzyme; dough sytrping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Secale cereale cultivar Halo.

PN WO200198474-A1.

PD 27-DEC-2001

PF 21-JUN-2001; 2001WO-BE00106

PR 22-JUN-2000; 2000GB-0015296
PR 25-JAN-2001; 2001GB-0002194
PR 26-JAN-2001; 2001GB-0002018
PR 16-MAR-2001; 2001GB-0006564
PR 21-MAY-2001; 2001GB-0012328

PA (LEUV-) LEUVEN RES & DEV

PI Delcour J, Debyser W, Gebruers K, Goesaert H, Fierens K, Robben J, Van Campenhout S;

DR WPI; 2002-114579/15.
DR N-PSDB; ABK13678.

PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or beta-glucanolytic enzymes comprises using endoxylanases during

PT screening for inhibition activity or affinity chromatography with
 XX immobilised enzymes -
 PS Claim 127; Page 60; 127pp; English.

The invention relates to separating and/or isolating inhibitors of cellulytic, xylanolytic and/or beta-glucanolytic enzymes comprises screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulase, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syruiping in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylanases such as by the formation of an endoxylanase/inhibitor complex. Improving the matting of cereals such as barley, sorghum and wheat and/or the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten starch separation and production, improving maize processing, plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence represents a partial rye SCX-I (S. cereale L endoxylanase inhibitor).

SQ Sequence 102 AA:

Query Match	56.8%	Score 105:	DB 23:	Length 102:
Best Local Similarity	87.0%	Pred. No. 4.1e-08:		
Matches 20: Conservative	1:	Mismatches 2:	Indels 0:	Gaps 0

```
QY      1 GAPVARAVEAVAFEGVCYDTKTL 23
          |||||:| |||||
Db      80 GAPVARAVKPVEPEGVCYDTKTL 102
```

Search completed: May 20, 2003, 16:14:37
Job time : 39.0769 secs

f:Accession: G86163
 f:Status: preliminary
 f:Molecule type: DNA
 A:Residues: 1-434 <SPO>
 A:Cross-references: GB:AE005172; NID:93850580; PIDN:AACT2120.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: conglutin gamma

Query Match 43.0%; Score 79.5; DB 2; Length 434;
 Best Local Similarity 47.1%; Pred. No. 0.002;
 Matches 16; Conservative 6; Mismatches 9; Indels 3; Gaps 1;

QY 5 ARAVEAVA---PGVCYDTKTLGNLGGYAVPNV 35
 DB 316 ARSIRKVASVKPGACFSTKNVGTSLGYAVPEI 349

RESULT 3

conglutin gamma precursor - narrow-leaved blue lupine
 C:Species: lupinus angustifolius (narrow-leaved blue lupine)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S21426; S31092
 R:Koliyas, S.; Gayler, K.R.
 Submitted to the EMBL Data Library, April 1992
 A:Reference number: S21426
 A:Accession: S21426
 A:Molecule type: mRNA
 A:Residues: 1-448 <KOL1>
 A:Cross-references: EMBL:X65601
 A:Note: It is uncertain whether Met-1 or Met-5 is the initiator
 R:Koliyas, S.; Gayler, K.R.
 Plant Mol. Biol. 21, 397-401, 1993
 A:Title: Structure of the cDNA coding for conglutin gamma, a sulphur-rich protein from L.
 A:Reference number: S31092; NMID:93144715; PMID:8425065
 A:Accession: S31092
 A:Molecule type: mRNA
 A:Residues: 5-448 <KOL2>
 A:Cross-references: EMBL:X65601
 C:Superfamily: conglutin gamma
 C:Keywords: :disulfide bond; seed; storage protein
 F:..33/Domain: signal sequence #status predicted <SIG>
 F:..34-294/Product: conglutin gamma large chain #status predicted <MA1>
 F:..95-448/Product: conglutin gamma small chain #status predicted <MA2>

Query Match 33.8%; Score 62.5; DB 1; Length 448;
 Best Local Similarity 42.4%; Pred. No. 0.57;
 Matches 14; Conservative 6; Mismatches 8; Indels 5; Gaps 2;

QY 3 PVARAVEAVAPFGVCYDTKTLGNLGGYAVPNV 35
 DB 335 PKQAVKAVGPGFGLCYDSRKIS---GG-APSV 362

RESULT 4

dermal glycoprotein precursor, extracellular - carrot (fragment)
 C:Species: Daucus carota (carrot)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Nov-2000
 C:Accession: T114329
 R:Stich, S.; Sturm, A.; Fujii, T.; Chrispeels, M.J.
 Planta 188, 432-438, 1992
 A:Title: cDNA cloning of an extracellular dermal glycoprotein of carrot and its expression
 A:Reference number: Z17987
 A:Accession: T114329
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-433 <SAT>
 A:Cross-references: EMBL:D14550; NID:d1004922; PID:d1003923
 C:Experimental source: subspecies US-Harumakigosun; root
 C:Genetics:
 A:Gene: EDGP1

C:Superfamily: conglutin gamma
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-433/Product: dermal glycoprotein, extracellular #status predicted <MAT>

Query Match 31.9%; Score 59; DB 2; Length 433;
 Best Local Similarity 39.3%; Pred. No. 1.7;
 Matches 11; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 8 VEAAPFGVCYDTKTLGNLGGYAVPNV 35
 DB 323 VASVAPPGACFSTDNILSTRLGPSVPSI 350

RESULT 5

GMP synthase VC0768 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: D82282
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: AB2035; NMID:20406833; PMID:10952301
 A:Accession: D82282
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-517 <HEI>
 A:Cross-references: GB:AE004162; GB:AE003852; NID:96655211; PIDN:AAF93933.1; GSPDB:GN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0768
 A:Map position: 1
 C:Superfamily: GMP synthase (glutamine-hydrolyzing); trpC homology

Query Match 30.5%; Score 56.5; DB 2; Length 517;
 Best Local Similarity 41.4%; Pred. No. 4.8;
 Matches 12; Conservative 2; Mismatches 6; Indels 9; Gaps 1;

QY 1 GAPVARAVEAVAPFGVCYDTKTLGNLGG 29
 DB 79 GVPV-----FGVCYGMOTMAEOLGG 98

RESULT 6

GMP synthase (glutamine-hydrolyzing) (EC 6.3.5.2) [validated] - Escherichia coli (str
 C:Species: Escherichia coli
 C:Date: 17-Mar-1987 #sequence_revision 30-Sep-1987 #text_change 01-Mar-2002
 C:Accession: A24640; B65027; A01197
 R:Tiedeman, A.A.; Smith, J.M.; Zalkin, H.
 J. Biol. Chem. 260, 8676-8679, 1985
 A:Title: Nucleotide sequence of the gua gene encoding GMP synthetase of Escherichia
 A:Reference number: A92946; NMID:85261223; PMID:3894545
 A:Accession: A24640
 A:Molecule type: DNA
 A:Residues: 1-525 <TIE>
 A:Cross-references: GB:M10101; GB:M10102; NID:g146274; PIDN:AA818619.1; PID:g146276
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; NMID:97426617; PMID:9278503
 A:Accession: B65027
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-525 <BLAT>
 A:Cross-references: GB:AE000337; GB:U00096; NID:q1788850; PIDN:AACT5560.1; PID:g17888
 A:Experimental source: strain K-12, substrain MG1655
 C:Comment: The active enzyme catalyzes the formation of GMP by transferring the amide
 C:Genetics:
 A:Gene: guaA

RESULT 11

Basic 7S globulin precursor - soybean

N:Alternate names: Insulin-binding protein

C:Species: Glycine max (soybean)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S06750; S00218; A61003; JN0311; JN0312

R:Kagawa, H.; Hirano, H.

Nucleic Acids Res. 17, 8868, 1989

A:Title: Sequence of a cDNA encoding soybean basic 7S globulin.

A:Reference number: S06750; M0ID:90067863; PMID:2587227

A:Accession: S06750

A:Molecule type: mRNA

A:Residues: 1-427 <KAI>

A:Cross-references: EMBL:X16469; NID:q18542; PIDN:CAA34489.1; PID:q18543

R:Kagawa, H.; Yamauchi, F.; Hirano, H.

FEBS Lett. 226, 145-149, 1987

A:Title: Soybean basic 7S globulin represents a protein widely distributed in legume spe

A:Reference number: S00218

A:Accession: S00218

A:Molecule type: protein

A:Residues: 276-293, 'X'; 295-304, 'FX'; 423-427, 'X'; 'IVGPFGLCPNNGVTSGLGPMXXMQPAR'; 'QL', 356,

42, 'AX' <KAG>

R:Hirano, H.; Matanabe, T.

Electrophoresis 11, 573-580, 1990

A:Title: Microsequencing of proteins electrophoretically transferred onto immobilizing matrices from

A:Reference number: A61003; M0ID:91031409; PMID:2226413

A:Accession: A61003

A:Molecule type: protein

A:Residues: 25-57; 135-150; 276-304, 'X'; 306-308, 'X'; 310-311; 383-417 <HIR>

R:Barbshov, S.F.; Egorov, T.A.; Kochkina, V.M.

Bioorg. Khim. 17, 421-423, 1991

A:Title: Isolation and characterization of soybean insulin-binding protein.

A:Reference number: JN0309; M0ID:91291181; PMID:2064630

A:Accession: JN0311

A:Molecule type: protein

A:Residues: 'VPVPOHKTN', 27-33, 'X'; 35-42, 'XX'; 45, 'X'; 47, 'X'; 49, 'AX'; 52 <BAR>

A:Note: 20-Thr, 21-Asn, 25-Pro and 50-Asn were also found

A:Accession: JN0312

A:Molecule type: protein

A:Residues: 'S', 99, 'IV', 102, 'SX', 105, 'C', 286-299, 'XX', 302, 'XX', 305 <BA2>

A:Note: 102-Gly and 296-Lys were also found.

C:Comment: This protein specifically binds to bovine insulin.

C:Superfamily: conglutin gamma

C:Keywords: heterodimer; seed; storage protein

3;1-24/Domain: signal sequence #status predicted <SIG>

3;23-27/Product: basic 7S globulin heavy chain #status experimental <MAT1>

1;216-427/Product: basic 7S globulin light chain #status experimental <MAT2>

Query Match

Best Local Similarity 29.2%; Score 54; DB 1; Length 427;

Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 2;

OY 3 PVARAVEAVAFEGVCYDPTKTLGNLGGYAVPNV 35

DB 314 PKQAVKSVAPFGLCFNS---NKINAY--PSV 340

RESULT 12

DNA-binding protein pabf - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 14-May-1999

C:Accession: T02029; T02342

R:Laidle, G.; Natt, E.; Seguin, A.; Doerner, P.W.; Lamb, C.J.

submitted to the EMBL Data Library, February 1994

A:Description: A novel DNA-binding protein with histone H1 and high mobility group prote

rometer.

A:Reference number: Z14501

A:Accession: T02029

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-546 <LAI>

A:Cross-references: EMBL:U06712; NID:9555654; PID:9555655

A:Experimental source: strain Xanthi; tissue-type stem

R:Tjden, G.; Coruzzi, G.M.

Plant Cell 1, 107-118, 1994

A:Title: A novel AT-rich DNA binding protein that combines a putative transcription

A:Reference number: Z14662

A:Accession: T02342

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 167-390, 'G', 392-546 <TJA>

A:Cross-references: EMBL:U26113; NID:9456123; PID:9456124

A:Experimental source: cultivar SRI; tissue-type leaf

C:Genetics:

A:Gene: pabf; ATPB-1

C:Keywords: DNA binding

Query Match

Best Local Similarity 29.2%; Score 54; DB 2; Length 546;

Matches 12; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

OY 1 GAPVARAVEAVAFEGVCYDPTKTLGNLGGYAVPN 34

DB 400 GPAAATVGVTDVPIAAFTENLPNAVGGGVYN 433

RESULT 13

dihydroorotate oxidase (EC 1.3.3.1) [imported] - Lactococcus lactis subsp. lactis (st

N:Alternate names: dihydroorotate dehydrogenase B

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: B86793

R:Bodlwin, A.; Muecker, P.; Manger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86623; M0ID:21235186; PMID:11337471

A:Accession: B86793

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <STO>

A:Cross-references: GB:AE005176; PID:q12724328; PIDN:AAK05444.1; GSPDB:GM00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: pydB

C:Superfamily: dihydroorotate oxidase

C:Keywords: oxidoreductase

Query Match

Best Local Similarity 28.9%; Score 53.5; DB 2; Length 311;

Matches 16; Conservative 5; Mismatches 12; Indels 13; Gaps 2;

OY 3 PVARAVEAVAFEGVCYDPTKTLGNLGGYAVPNV 35

DB 180 PIRAAVEAGACDGLTINTLMGVRFDLTKRPVLANITGSLSPAI 225

RESULT 14

GMP synthase PA3769 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83173

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; M0ID:20437337; PMID:10984043

A:Accession: G83173

A:Status: preliminary

A:Molecule type: DNA

THIS PAGE BLANK (use 70)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 : Search time 9.42308 Seconds
(without alignments)
154.055 million cell updates/sec

Title: US-09-869-155-13

Perfect score: 185
Sequence: 1 GAPVRAVEAVAFEGVCYDTKLTGNNLCGYAVPNV 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56.5	30.5	517	1	GUAA_VIBCH
2	56.5	30.5	525	1	GUAA_ECOLI
3	54	29.2	427	1	7SBG_SOYBN
4	53.5	28.9	311	1	PYDB_IACLA
5	53.5	28.9	311	1	PYDB_IACLC
6	52.5	28.4	525	1	GUAA_PSEAE
7	52	28.1	212	1	AGI_HORVU
8	52	28.1	3206	1	POLG_PSBMV
9	51	27.6	556	1	CD19_HUMAN
10	51	27.6	944	1	SYL2_SULSO
11	50.5	27.3	500	1	DHAL_AGARI
12	50.5	27.3	525	1	GUAA_AGR75
13	50	27.0	430	1	CD19_CAVPO
14	50	27.0	432	1	YK27_CAEEL
15	50	26.8	534	1	SR54_ASPNG
16	49.5	26.8	510	1	GUAA_AQUAE
17	49.5	26.8	522	1	GUAA_XYLFA
18	48.5	26.2	255	1	AMPR_TREPA
19	48.5	26.2	524	1	GUAA_CORAM
20	48.5	26.2	526	1	GUAA_STRCO
21	48	25.9	186	1	AG13_WHEAT
22	48	25.9	213	1	AG12_WHEAT
23	48	25.9	901	1	MSH5_YEAST
24	48	25.9	1004	1	SLPO_BACBR
25	48	25.9	1608	1	HLTA_SERMA
26	48	25.9	2185	1	PYR1_DICDI
27	47.5	25.7	335	1	GPDA_PSELE
28	47.5	25.7	523	1	GUAA_HAEIN
29	47	25.4	231	1	YC65_ARCFU
30	47	25.4	301	1	MCP_BPR41
31	47	25.4	432	1	PROA_DEIRA
32	47	25.4	477	1	NIFE_RHOCA
33	47	25.4	547	1	CD19_MOUSE

34	47	25.4	675	1	CAAA_BACTI
35	47	25.4	915	1	SMOO_HUMAN
36	46.5	25.1	521	1	GUAA_NEIMB
37	46.5	25.1	589	1	GLMS_METTH
38	46.5	25.1	782	1	DPO2_ECOLI
39	46.5	25.1	4485	1	DYHG_CHLRE
40	46	24.9	269	1	MURI_PASMT
41	46	24.9	430	1	GSA_BACSU
42	46	24.9	432	1	GSA_VIBCH
43	46	24.9	507	1	GUAA_BRUME
44	46	24.9	571	1	PRDE_HUMAN
45	46	24.9	614	1	BTUH_ECOLI

ALIGNMENTS

RESULT 1	GUAA_VIBCH	STANDARD:	PRT:	517 AA.
ID	GUAA_VIBCH			
AC	O9KTW2:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (Glutamine			
DE	amidotransferase) (GMP synthetase).			
GN	GUAA OR VC0768.			
OS	Vibrio cholerae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.			
OX	NCBI_TaxID=666;			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=El Tor N16961 / Serotype O1;			
RC	MEDLINE=20406833; PubMed=10952301;			
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,			
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,			
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,			
RA	Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,			
RA	McDonald L., Uitterlind T., Fleischmann R.D., Niernan W.C., White O.,			
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,			
RA	Fraser C.M.;			
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio			
RT	cholerae.";			
RL	Nature 406:477-483(2000).			
CC	-1- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +			
CC	H(2O) -> AMP + diphosphate + GMP + L-glutamate.			
CC	-1- PATHWAY: GMP biosynthesis.			
CC	-1- SUBUNIT: HOMODIMER (BY SIMILARITY).			
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMP SYNTHASE			
CC	FAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL: AE004162; AAF93933.1; -			
CC	HSSP: P04079; IGPM.			
DR	TIGR: VC0768; -			
DR	InterPro: IPR002385; Anth_synthII.			
DR	InterPro: IPR001317; CPS_GATase.			
DR	InterPro: IPR000991; GATase_1.			
DR	InterPro: IPR001674; GMP_synth_C.			
DR	InterPro: IPR004739; GMPsynthase_N.			
DR	Pfam: PF00917; GATase_1.			
DR	Pfam: PF00958; GMP_synth_C_1.			
DR	PRINTS: PR00097; ANTSYNTHASEII.			
DR	PRINTS: PR00099; CPSGATASE.			
DR	PRINTS: PR00096; GATASE.			

DR TIGR00884; guaa_Cterm; 1.
 DR TIGR00888; guaa_Nterm; 1.
 DR PROSITE; PS00442; GATASE_Type_1; 1.
 KM Ligase: GMP biosynthesis: Purine biosynthesis: ATP-binding;
 KM Glutamine amidotransferase; Complete proteome;
 FT DOMAIN 1 198
 FT ACT_SITE 231 393 GMP-BINDING (BY SIMILARITY).
 FT ACT_SITE 86 86 GATASE (BY SIMILARITY).
 FT ACT_SITE 173 173 GATASE (BY SIMILARITY).
 FT ACT_SITE 175 175 GATASE (BY SIMILARITY).
 FT NP_BIND 227 233 ATP (BY SIMILARITY).
 SQ SEQUENCE 517 AA; 57764 MW; 80AC2194E1172C45 CRC64;

Query Match 30.5%; Score 56.5; DB 1; Length 517;
 Best Local Similarity 41.4%; Pred. No. 3.1;
 Matches 12; Conservative 2; Mismatches 6; Indels 9; Gaps 1;

QY 1 GAPYARAEVAPGVCYDRTKGNLNG 29
 Db 79 GVPV-----FGVCYGMQTMALQGH 98

RESULT 2
 GUA_ECOLI STANDARD; PRT; 525 AA.
 ID GUA_ECOLI
 AC P04079;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (glutamine
 DE amidotransferase) (GMP synthetase).
 GN GUA OR B2507.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 FN (1)
 FP SEQUENCE FROM N.A.
 FX MEDLINE=8526123; PubMed=3894345;
 RA Tiedeman A.A., Smith J.M., Zalkin H.;
 RT "nucleotide sequence of the guaa gene encoding GMP synthetase of
 RT Escherichia coli K12";
 RL J. Biol. Chem. 260:8676-8679(1985).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
 RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horinouchi T.;
 RT "Construction of a contiguous 874-Kb sequence of the Escherichia coli
 RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
 RT analysis of its sequence features";
 RL DNA Res. 4:91-113(1997).
 RN (4)
 RP GATASE DOMAIN.
 RX MEDLINE=85131126; PubMed=2982857;
 RA Zalkin H., Argos P., Narayana S.V.L., Tiedeman A.A., Smith J.M.;
 RT "Identification of a trpG-related glutamine amide transfer domain in
 RT Escherichia coli GMP synthetase.";

RL J. Biol. Chem. 260:3350-3354(1985).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=96133732; PubMed=8548458;
 RA Tesmer J.J.G., Klem T.J., Deras M.L., Davison V.J., Smith J.L.;
 RT "The crystal structure of GMP synthetase reveals a novel catalytic
 RT triad and is a structural paradigm for two enzyme families.";
 RL Nat. Struct. Biol. 3:74-86(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
 CC H(2)O = AMP + diphosphate + GMP + L-glutamate.
 CC -1- PATHWAY: GMP biosynthesis.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMP SYNTHASE
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M10101; AAB18619.1; -
 CC DR EMBL: AE000337; AAC75560.1; -
 CC DR EMBL: D90880; BAA16394.1; -
 CC DR PIR: A24640; SYECGU.
 CC DR PDB: 1GPM; 29-JAN-96.
 CC DR Ecogene: EG10420; guaa.
 CC DR Interpro: IPR000991; GATase_1.
 CC DR Interpro: IPR001674; GMP_synth_C.
 CC DR Interpro: IPR004739; GMPsynthase_N.
 CC DR Interpro: IPR004135; tRNA_Me_trans.
 CC DR Pfam: PF00117; GATase; 1.
 CC DR Pfam: PF00958; GMP_synth_C; 1.
 CC DR Pfam: PF03054; tRNA_Me_trans; 1.
 CC DR PRINTS: PR00096; GATASE.
 CC DR TIGR00884; guaa_Cterm; 1.
 CC DR TIGR00888; guaa_Nterm; 1.
 CC DR PROSITE: PS00442; GATASE_Type_1; 1.
 CC KW Ligase: GMP biosynthesis; Purine biosynthesis; ATP-binding;
 CC KW Glutamine amidotransferase; 3D-structure; Complete proteome.
 CC FT DOMAIN 1 206
 CC FT ACT_SITE 239 401 GMP-BINDING.
 CC FT ACT_SITE 86 86 GATASE.
 CC FT ACT_SITE 181 181 GATASE.
 CC FT NP_BIND 235 241 ATP.
 CC SQ SEQUENCE 525 AA; 58679 MW; D934786DF3D694B CRC64;

Query Match 30.5%; Score 56.5; DB 1; Length 525;
 Best Local Similarity 52.2%; Pred. No. 3.2;
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 9 EAVAP-FGVCYDRTKGNLNGY 30
 Db 77 EACVPFVCYGMQTMALQGH 99

RESULT 3
 TSBG_SOYBN STANDARD; PRT; 427 AA.
 ID TSBG_SOYBN
 AC P13917; O39901; Q43464;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Basic 7S globulin precursor (Bg) (SBg7S).
 GN BG OR G7S.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

```

OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Miyagishirume; TISSUE=Seed;
RX MEDLINE=90067863; PubMed=2587227;
RA Kagawa H., Hirano H.;
RT "Sequence of a cDNA encoding soybean basic 7S globulin.";
RN Nucleic Acids Res. 17:8868-8868(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Miyagishirume; TISSUE=Etolated leaf;
RX MEDLINE=94336768; PubMed=8058830;
RA Watanabe Y., Hirano H.;
RT "Nucleotide sequence of the basic 7S globulin gene from soybean.";
RN Plant Physiol. 105:1019-1020(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Shi-shi; TISSUE=Colyledon;
RX Shu T.F., Hsieh K.L., Hsing Y.I., Chen Z.Y., Chow T.Y.;
RT Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE.
RA Kagawa H., Yamauchi F., Hirano H.;
RT "Soybean basic 7S globulin represents a protein widely distributed in
legume species.";
RN FEBS Lett. 226:145-149(1987).
CC -1- FUNCTION: SEED STORAGE PROTEIN. HAS A PROTEIN KINASE ACTIVITY.
CC BINDS LEGUMININ.
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBUNIT: THE MATURE BG IS COMPOSED OF HIGH- AND LOW-KDA SUBUNITS
WHICH ARE LINKED BY DISULFIDE BONDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X16469; CAA34489.1; -
DR EMBL: U59425; AAR03390.1; -
DR EMBL: D16107; BAA03681.1; -
DR PIR: S06750; S06750.
KW Glycoprotein; Seed storage protein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 427 BASIC 7S GLOBULIN.
FT CHAIN 25 275 HIGH-KDA SUBUNIT.
FT CHAIN 276 427 LOW-KDA SUBUNIT.
FT CONFLICT 48 48 W->S (IN REF. 3).
FT CONFLICT 264 264 N->T (IN REF. 1).
FT CONFLICT 305 305 F->C (IN REF. 1).
FT CONFLICT 309 309 F->C (IN REF. 1).
SQ SEQUENCE 427 AA; 46393 MW; 66041BC0680BACB CRC64;

Query Match 29.2%; Score 54; DB 1; Length 427;
Best Local Similarity 36.4%; Pred. No. 5.8;
Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 2;

QY 3 PVARAVEAVAPGVCYDPTLTGNNLCGYAVPNV 35
DB 314 PROAQKVSAPGCLFENS---NKINAV--PSV 340

RESULT 4
PYDB_LACLA STANDARD; PRT; 311 AA.
AC Q9CFW8;
DR 16-OCT-2001 (Rel. 40; Created)
DR 16-OCT-2001 (Rel. 40; Last sequence update)
DR 15-JUN-2002 (Rel. 41; Last annotation update)
DE Dihydroorotate dehydrogenase B, catalytic subunit (EC 1.3.3.1)
DE (Dihydroorotate oxidase B) (DHODase B) (DHOD B).

```

```

GN PYRDB OR PYDB OR LL1346.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RT Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403.";
RN Genome Res. 11:731-753(2001).
CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
H(2O(2)).
CC -1- COFACTOR: FMN (By similarity).
CC -1- PATHWAY: Pyrimidine biosynthesis; fourth step.
CC -1- SUBUNIT: HETEROETRAMER OF 2 PYR AND 2 PYRDB SUBUNITS (By
similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DIHYDROOROTATE DEHYDROGENASE FAMILY.
SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE006366; AAK05444.1; -
DR HSSP: P54321; 2DOR.
DR InterPro: IPR001295; DHO_dh.
DR InterPro: IPR003009; FMN_enzyme.
DR Pfam: PF01180; DHODase; 1.
DR TIGRfam: TIGR01037; pyrd_sub1_fam; 1.
DR PROSITE: PS00911; DHODENASE_1; 1.
DR PROSITE: PS00912; DHODENASE_2; 1.
KW Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN;
Complete proteome.
FT NP_BIND 242 250 FMN (POTENTIAL).
FT SEQUENCE 311 AA; 32916 MW; 6FE419606B2E733C CRC64;

Query Match 28.9%; Score 53.5; DB 1; Length 311;
Best Local Similarity 34.8%; Pred. No. 5;
Matches 16; Conservative 5; Mismatches 12; Indels 13; Gaps 2;

QY 3 PVARAVEAVAP-----FGVCYDPTKT---LGNLCGYAVPNV 35
DB 180 PIKAVEAGAGDGLTINTLMGVRFPLKRPVLANITGSLSPAI 225

RESULT 5
PYDB_LACLC STANDARD; PRT; 311 AA.
AC P54322;
DR 01-OCT-1996 (Rel. 34; Created)
DR 01-OCT-1996 (Rel. 34; Last sequence update)
DR 15-JUN-2002 (Rel. 41; Last annotation update)
DE Dihydroorotate dehydrogenase B, catalytic subunit (EC 1.3.3.1)
DE (Dihydroorotate oxidase B) (DHODase B) (DHOD B).
GN PYRDB.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1363;
RX MEDLINE=94292416; PubMed=8021180;
RA Andersen P.S., Jansen P.J.G., Hammer K.;
RT "Two different dihydroorotate dehydrogenases in Lactococcus lactis.";
RN J. Bacteriol. 176:3975-3982(1994).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MG1363;
RX MEDLINE-96326348; PubMed-8759867;
RA Andersen P.S., Martinussen J., Hammer K.;
RT "Sequence analysis and identification of the pyrKDF operon from
RT Lactococcus lactis including a novel gene, pyrK, involved in
RT pyrimidine biosynthesis."
RL J. Bacteriol. 178:5005-5012(1996).
RN [3]
RP CHARACTERIZATION.
RC STRAIN-MG1363;
RX MEDLINE-97067197; PubMed-8910599;
RA Nielsen F.S., Andersen P.S., Jensen K.F.;
RT "The B form of dihydroorotate dehydrogenase from Lactococcus lactis
RT consists of two different subunits, encoded by the pyrB and pyrK
RT genes, and contains FMN, FAD, and [Fes] redox centers."
RL J. Biol. Chem. 271:29359-29365(1996).
CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
CC H(2)O(2).
CC -1- COFACTOR: FMN.
CC -1- PATHWAY: pyrimidine biosynthesis; fourth step.
CC -1- SUBUNIT: HETEROTRIMER OF 2 PYR AND 2 PYRDB SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -1- MISCELLANEOUS: CAN USE NAD BUT NOT FMN AS AN ELECTRON
CC ACCEPTOR.
CC -1- SIMILARITY: BELONGS TO THE DIHYDROOROTATE DEHYDROGENASE FAMILY.
CC SUBFAMILY 1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X74207; CAA52280.1; -
DR HSSP: P54321; 2DOR.
DR InterPro: IPR001295; DHO_dh.
DR InterPro: IPR003009; FMN_enzyme.
DR Pfam: PF01180; DHODEHase_1.
DR TIGRPFAM: TIGR01037; pyrD_sub1_fam.1.
DR PROSITE: PS00911; DHODEHase_1; 1.
DR PROSITE: PS00912; DHODEHase_2; 1.
DR Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN.
FT NP_BIND 242 250 FMN (POTENTIAL).
FT SEQUENCE 311 AA; 33100 MW; E64C5B635F57B6EE CRC64;
SQ
Query Match 28.9%; Score 53.5; DB 1; Length 311;
Best Local Similarity 34.8%; Pred. No. 5;
Matches 16; Conservative 5; Mismatches 12; Indels 13; Gaps 2;
OY 3 PVRBAEAVAP-----FGVCYDRTK-----LGNNGSYAVPNV 35
Db 180 PIKAAVEAAGADGTLNIMLMGVFDPKTRQPILANITGLSPAI 225
RESULT 6
GUA_PSEAE
ID GUA_PSEAE STANDARD: PRT; 525 AA.
AC Q9HXM6;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (Glutamine
DE amidotransferase) (GMP synthetase).
GN GUA OR PA3769.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
CC H(2)O = AMP + diphosphate + GMP + L-glutamate.
CC -1- PATHWAY: GMP biosynthesis.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMP SYNTHASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE004796; AAC07156.1; -
DR HSSP: P04079; 1GPM.
DR InterPro: IPR002385; Anth_synthII.
DR InterPro: IPR001317; CPS_GATase.
DR InterPro: IPR000991; GATase_1.
DR InterPro: IPR001674; GMP_synth_C.
DR InterPro: IPR004739; GMPsynthase_N.
DR Pfam: PF00117; GATase_1.
DR Pfam: PF00958; GMP_synth_C; 1.
DR PRINTS: PR00097; ANTSNTASEII.
DR PRINTS: PR00099; CPSGATASE.
DR PRINTS: PR00096; GATASE.
DR TIGRPFAM: TIGR00884; guaA_cleam.1.
DR TIGRPFAM: TIGR00888; guaA_nterm.1.
DR PROSITE: PS00442; GATASE_TYPE1; 1.
DR Ligase; GMP biosynthesis; Purine biosynthesis; ATP-binding;
KW Glutamine amidotransferase; Complete proteome.
FT DOMAIN 1 206
FT ACT_SITE 239 401 GMP-BINDING (BY SIMILARITY).
FT ACT_SITE 86 86 GATASE (BY SIMILARITY).
FT ACT_SITE 181 181 GATASE (BY SIMILARITY).
FT ACT_SITE 183 183 GATASE (BY SIMILARITY).
FT NP_BIND 235 241 ATP (BY SIMILARITY).
FT SEQUENCE 525 AA; 57959 MW; CDBE1CCE73C3C40 CRC64;
SQ
Query Match 28.4%; Score 52.5; DB 1; Length 525;
Best Local Similarity 41.4%; Pred. No. 11;
Matches 12; Conservative 5; Mismatches 7; Indels 5; Gaps 2;
OY 6 RAVEAV-----AP-FGVCYDRTKGNLGG 29
Db 70 RAQNAVFDLKVPLFGICYGQMTAEQMG 98
RESULT 7
AGI_HORVU
ID AGI_HORVU STANDARD: PRT; 212 AA.
AC P15312;
DT 01-APR-1990 (Rel. 14; Created)
DT 01-APR-1990 (Rel. 14; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Root-specific lectin precursor.
DE Hordeum vulgare (Barley).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;

```


OC Trilicaceae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lerner D.R., Kaihel N.V.;
 RT "Cloning and characterization of root-specific barley lectin.";
 RL Plant Physiol. 91:124-129(1989).
 CC -1- FUNCTION: CARBOHYDRATE BINDING.
 CC -1- SIMILARITY: CONTAINS 4 CHITIN-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29280; AAA32969.1; .
 DR HSSP; P10969; 1MGT.
 DR InterPro: IPR001002; Chitin_binding_1.
 DR Pfam: PF00187; Chitin_bind_1; 4.
 DR PRINTS; PR00451; CHITINBINDING.
 DR ProDom; PD000609; Chitin_binding_1; 4.
 DR SMART; SM00270; ChbD1; 4.
 DR PROSITE; PS00026; CHITIN_BINDING; 4.
 DR Lerner; Repeat: Chitin-binding; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 212
 FT DOMAIN 27 69 ROOT-SPECIFIC LECTIN.
 FT DOMAIN 70 112 CHITIN-BINDING 1.
 FT DOMAIN 113 155 CHITIN-BINDING 2.
 FT DOMAIN 156 197 CHITIN-BINDING 3.
 FT MOD_RES 27 27 PYRROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT DISULFID 29 44 BY SIMILARITY.
 FT DISULFID 38 50 BY SIMILARITY.
 FT DISULFID 43 57 BY SIMILARITY.
 FT DISULFID 61 66 BY SIMILARITY.
 FT DISULFID 72 87 BY SIMILARITY.
 FT DISULFID 81 93 BY SIMILARITY.
 FT DISULFID 86 100 BY SIMILARITY.
 FT DISULFID 104 109 BY SIMILARITY.
 FT DISULFID 115 130 BY SIMILARITY.
 FT DISULFID 124 136 BY SIMILARITY.
 FT DISULFID 129 143 BY SIMILARITY.
 FT DISULFID 147 152 BY SIMILARITY.
 FT DISULFID 158 173 BY SIMILARITY.
 FT DISULFID 167 179 BY SIMILARITY.
 FT DISULFID 172 186 BY SIMILARITY.
 FT DISULFID 190 195 BY SIMILARITY.
 FT CARBOHYD 206 206 N-LINKED (GLCNAC...).
 SQ SEQUENCE 212 AA; 21209 MW; 8D948245DB6B25A5 CRC64;
 Query Match 28.1%; Score 52; DB 1; Length 212;
 Best Local Similarity 45.0%; Pred. No. 5.6;
 Matches 9; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

DE (6R2): Genome-linked protein (VPG): Nuclear inclusion protein A (NI-A)
 DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro): Nuclear
 DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC
 DE 2.7.7.48): Coat protein (CP)].
 OS Pea seed-borne mosaic virus (strain DPDI).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=31736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92044431; PubMed=1940858;
 RA Johansen E., Rasmussen O.F., Heide M., Borkhardt B.;
 RT "The complete nucleotide sequence of pea seed-borne mosaic virus
 RNA.";
 RL J. Gen. Virol. 72:2625-2632(1991).
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
 CC further restricted by preferences for the amino acids in p6 - p1'
 CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 CC Glu+(Ser or Gly) for the enzyme from tobacco etch virus. The
 CC natural substrate is the viral polyprotein, but other proteins and
 CC oligopeptides containing the appropriate consensus sequence are
 CC also cleaved.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-I-Gly bond at its own C-
 CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-I-Gly, in the
 CC processing of the polyviral polyprotein.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10930; BAA01726.1; .
 DR PIR; J01331; GNVSFV.
 DR MEROPS; C04.010; .
 DR MEROPS; C06.001; .
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001456; Peptidase_C6.
 DR InterPro: IPR002540; Poly_P1.
 DR InterPro: IPR001592; Poly-coat.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00767; Poly-coat; 1.
 DR Pfam; PF00851; Peptidase_C6; 1.
 DR Pfam; PF00863; Peptidase_C4; 1.
 DR Pfam; PF01577; Poly_P1; 1.
 DR PRINTS; PR00966; NIAPOLYPTASE.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.
 DR Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 ATP-binding.

[illegible]

```

RL Immunogenetics 35:102-111(1992).
CC -1- FUNCTION: MAY BE INVOLVED IN GROWTH REGULATION OF B-CELLS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 396
CC ONWARD AND IS SHORTER (467 AA) DUE TO A FRAMESHIFT.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide cd19 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd19.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL, M28170; AAA68490.1; -
CC DR EMBL, M21097; AAA35533.1; ALT_FRAME.
CC DR EMBL, M84371; AAA69966.1; -
CC DR EMBL, M62550; AAB60697.1; -
CC DR EMBL, M62544; AAB60697.1; JOINED.
CC DR EMBL, M62545; AAB60697.1; JOINED.
CC DR EMBL, M62546; AAB60697.1; JOINED.
CC DR EMBL, M62547; AAB60697.1; JOINED.
CC DR EMBL, M62548; AAB60697.1; JOINED.
CC DR EMBL, M62549; AAB60697.1; JOINED.
CC DR PIR, J10074; J10074.
CC DR PIR, A45808; A45808.
CC DR PIR, A44441; A44441.
CC DR GENE, HGNC:1633; CD19.
CC DR MIM, 107265;
CC DR InterPro: IPR003599; Iq.
CC DR InterPro: IPR003006; Iq_MHC.
CC DR Pfam: PF00047; Iq; 2.
CC DR SMART, SM00409; Iq; 2.
CC KW Immunoglobulin domain; Glycoprotein; Transmembrane; B-cell; Signal;
CC Repeat.
CC FT SIGNAL 1 19 POTENTIAL.
CC FT CHAIN 20 556 B-LYMPHOCYTE ANTIGEN CD19.
CC FT DOMAIN 20 291 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 292 313 POTENTIAL.
CC FT DOMAIN 314 556 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 31 104 Iq-LIKE C2-TYPE DOMAIN 1.
CC FT DOMAIN 193 268 Iq-LIKE C2-TYPE DOMAIN 2.
CC FT DISULFID 38 97 POTENTIAL.
CC FT CARBOHYD 200 261 POTENTIAL.
CC FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 29 29 E -> EG (IN REF. 4).
CC FT CONFLICT 80 80 S -> I (IN REF. 2).
CC FT CONFLICT 186 186 O -> QAEIVSLPVP (IN REF. 3).
CC FT CONFLICT 514 514 H -> R (IN REF. 3 AND 4).
CC FT CONFLICT 556 556 AA; 61069 MW; 5957952A169AFB42 CRC64;
CC SO SEQUENCE
CC -----
CC Query Match 27.6%; Score 51; DB 1; Length 556;
CC Best Local Similarity 27.9%; Pred. No. 19;
CC Matches 12; Conservative 7; Mismatches 14; Indels 10; Gaps 1,
CC
CC 3 PVARAVEAVAPGVCYDTKTLGNINIGG-----YAVPV 35
CC 1111:::1111:1111:1111:
CC Db 471 PVARTMDPLSPHGSAMDPSREATSLGSSQSYEDMRCILYLAAPOL 513
CC
CC RESULT 10
CC SYL2_SULSO STANDARD. PRT. 944 AA.
CC ID SYL2_SULSO
CC AC 033768; 090W23;
CC 15-DEC-1998 (Rel. 37, Created)

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE leucyl-tRNA synthetase 2 (EC 6.1.1.4) (Leucine--tRNA ligase 2) (LeuRS
 2)
 GN LEU52 OR LEU5 OR SSO0589 OR C21_009.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis B., Duguet M., Erasuo G., Faguy D.,
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kushiwa N., Lafleur E., Medina N., Peng X., Penny S.L., She O.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A., Sensen C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awezaz M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Helamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RN [3]
 RP SEQUENCE OF 762-944 FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=97352708; PubMed=9209067;
 RA Charlebois R.L., Sensen C.W., Doolittle W.F., Brown J.R.;
 RT "Evolutionary analysis of the hscGABRFDH1 gene cluster from the
 archaeon Sulfolobus solfataricus P2.";
 RL J. Bacteriol. 179:4429-4432(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
 CC diphosphate + L-leucyl-tRNA(Leu).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y18930; CAB57710.1; -
 DR EMBL: AE006689; AAK40902.1; -
 DR EMBL: U82227; AAB63016.1; -
 DR InterPro: IPR004493; Leu-tRNA-synt_1a.
 DR InterPro: IPR002300; tRNA-synt_1a.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00133; tRNA-synt_1; 1.
 DR TIGRFAMs: TIGR00395; leuS_arch; 1.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_I; FALSE_NEG.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 KW SITE 36 46 "HIGH" REGION.
 FT SITE 623 627 "KMSKS" REGION.
 FT BINDING 626 626 ATP (BY SIMILARITY).
 SQ SEQUENCE 944 AA; 109104 MW; 9ACAPCFBFO4160DC CRC64;

Query Match

27.6%; Score 51; DB 1; Length 944;

Best Local Similarity 46.2%; Pred. No. 31;
 Matches 12; Conservative 1; Mismatches 13; Indels 0; Gaps 0;
 QY 1 GAPVAAVEAVAPFGVCYDTKLGNN 26
 DB 301 GTGVAAVSPAHAPFDYYSKTEKKN 326
 RESULT 11
 ID DHAL_AGABI STANDARD; PRT; 500 AA.
 AC 074187;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aldenhyde dehydrogenase (EC 1.2.1.3) (ALDH).
 GN ALDA.
 OS Agaricus bisporus (Common mushroom).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Agaricaceae; Agaricus.
 NCBI_TaxID=5341;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Horst H39;
 RA Schapp P.J., Muller Y., Visser J.;
 RT "Molecular structure and spatial expression of housekeeping genes in
 RT mushrooms";
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2O) = an acid + NADH.
 CC -1- PATHWAY: Ethanol utilization; second step.
 CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y17825; CAA76875.1; -
 DR HSSP: P51977; IBXS.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR Pfam: PF00171; aldehyd; 1.
 DR PROSITE: PS006070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Oxidoreductase; NAD.
 FT NP_BIND 246 251 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 269 269 BY SIMILARITY.
 FT ACT_SITE 303 303 BY SIMILARITY.
 SQ SEQUENCE 500 AA; 54395 MW; C4FCE5B50855925 CRC64;
 Query Match 27.3%; Score 50.5; DB 1; Length 500;
 Best Local Similarity 40.0%; Pred. No. 20;
 Matches 12; Conservative 4; Mismatches 13; Indels 1; Gaps 1;
 QY 5 ARAVEAVAPFGVCYDTKLGNNLGYAVPN 34
 DB 457 ANTIDAGVPPG-GYKQSGIGRELGAYALNN 485
 RESULT 12
 ID GUAA_AGRT5 STANDARD; PRT; 525 AA.
 AC 080112;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (Glutamine
 DE amidotransferase) (GMP synthetase).
 GN GUA OR ATU0281 OR AGR_C_480.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

```

OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Seubai J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Glendinning J., Deatherage G., Gillet W., Grant C.,
RA Kuyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Senphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RT Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Gourollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Hummel K., Gordon J., Vaudin M., Ratchouk O., Epp A., Liu F.,
RA Wcliam C., Allinger M., Doughy D., Scott C., Lapps C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Chelo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RT Science 294:2323-2328(2001).
CC - - CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
CC H(2)O -> AMP + diphosphate + GMP + L-glutamate.
CC - - PATHWAY: GMP biosynthesis.
CC - - SUBUNIT: Homodimer (by similarity).
CC - - SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMP SYNTHASE
CC FAMILY.
CC - - SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/).
CC cr send an email to license@sib-sib.ch).
CC -----
DR EMBL; AE009000; AAL1303.1; -
DR EMBL; AE007967; AAK86096.1; -
DR PROSITE; PS00442; GATASE_TYPE_1; 1.
KW lIase; GMP biosynthesis; Purine biosynthesis; ATP-binding;
KM Glutamine amidotransferase; Complete proteome.
FT FT DOMAIN 1 201
FT NP_BIND 230 236
FT FT 234 401
FT ACT_SITE 86 86
FT ACT_SITE 176 176
FT ACT_SITE 178 178
SQ SEQUENCE 525 AA; 57521 MW; 581CD3C31F64F34 CRC64;
Query Match 27.3%; Score 50.5; DB 1; Length 525;
Best Local Similarity 44.1%; Pred. NO. 21;
Matches 15; Conservative 4; Mismatches 8; Indels 7; Gaps 3;
QY 1 GAPVARAVEAV---AP-FGVGYDKRTGNNNGG 29
Db 67 GSP--RAPQAVFDSKVPILGICYGQDTLCTOLGG 98

```

```

DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen CD19 (B-lymphocyte surface antigen B4) (LEU-12)
DE (Differentiation antigen CD19) (Fragment).
GN CD19.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognath; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91332460; PubMed=1714482;
RA Zhou L.J., Ord D.C., Hughes A.L., Tedder T.F.;
RT "Structure and domain organization of the CD19 antigen of human,
RT mouse, and guinea pig B lymphocytes. Conservation of the extensive
RT cytoplasmic domain."
RL J. Immunol. 147:1424-1432(1991).
CC - FUNCTION: MAY BE INVOLVED IN GROWTH REGULATION OF B-CELLS.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: BELONGS TO THE IMMUGLOBULIN SUPERFAMILY.
CC - SIMILARITY: CONTAINS 2 IMMUGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M62543; AAA51529.1; -.
DR InterPro; IPR003399; 1g_-.
DR InterPro; IPR003006; 1g_MHC.
DR Pfam; PF00047; 1g; 1.
DR SMART; SM00409; 1G; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; B-cell; Repeat.
FT NON_TER 1
FT DOMAIN 1 168 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 169 190 POTENTIAL.
FT DOMAIN 191 430 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 431 17 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 69 145 IG-LIKE C2-TYPE DOMAIN 2.
FT DISULF 76 138 POTENTIAL.
FT CARBOHYD 1 1 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 430 AA; 46957 MW; 84F05B04F94590E3 CRC64;

Query Match 27.0%; Score 50; DB 1; Length 430;
Best Local Similarity 27.9%; Pred. No. 20;
Matches 12; Conservative 6; Mismatches 15; Indels 10; Gaps 1;

3 PVARAVEAVAFPGVCYDRTKLTGNNIGG-----YAVPNV 35
||| : : : : : : : ||
Db 347 PVARTEDFLSPHGSAMDPSEATSLGSGSYEDMRCILVAAPQL 389

RESULT 14
YK27_CAEEL STANDARD; PRT; 432 AA.
AC 016686;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 50.0 kDa protein K07E8.7 in chromosome IV.
GN K07E8.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitida; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
NN [1]

```

RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Jones K., Kramer J.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE RLU FAMILY OF PSEUDOURIDINE SYNTHASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF016678; AAB6153.1;
 CC Wormpep: K07E8.7; CELL1898.
 DR InterPro: IPR002990; PSL_RLU.
 DR InterPro: IPR000613; Pseudou_synth.
 DR Pfam: PF00849; Pseudou_synth_2; 1.
 DR TIGRFAMs: TIGR00005; rlud_subfam; 1.
 DR PROSITE: PS01129; PSL_RLU; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 432 AA; 50011 MW; ABC9084ABA84734 CRC64;

Query Match 27.0%; Score 50; DB 1; Length 432;
 Best Local Similarity 36.0%; Pred. No. 20;
 Matches 9; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 4 VARAEVAPFGVCYDTKTLGNLNG 28
 DB 60 IDGVNRHLAPYMACYRTRKGRWIG 84

RESULT 15
 SR54.ASPNG
 ID SR54.ASPNG STANDARD; PRT; 534 AA.
 AC 000179;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Signal recognition particle 54 kDa protein homolog.
 GN SRPA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96144300; PubMed=8566805;
 RA Thompson S.A., Gollightly E.J., Yaver D.S.;
 RT "Nucleotide sequence of the Aspergillus niger srpa gene."
 RL Gene 167:337-338(1995).
 CC -1- FUNCTION: BINDS TO THE SIGNAL SEQUENCE OF PRESECRETORY PROTEIN
 CC WHEN THEY EMERGE FROM THE RIBOSOMES AND TRANSFERS THEM TO TRAM
 CC (TRANSLOCATING CHAIN-ASSOCIATING MEMBRANE PROTEIN).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DOMAIN: HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN BINDS GTP; THE
 CC M-DOMAIN BINDS THE 7S RNA IN PRESENCE OF SRP19 AND ALSO BINDS THE
 CC SIGNAL SEQUENCE.
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L38317; AAB04946.1;
 DR HSSP: O07347; 2FPH.
 DR InterPro: IPR000897; SRP54.
 DR InterPro: IPR004125; SRP54_SPB.

DR Pfam: PF00448; SRP54; 1.
 DR Pfam: PF02881; SRP54_N; 1.
 DR Pfam: PF02978; SRP_SPB; 1.
 DR PRODOM: PD000819; SRP54; 1.
 DR PROSITE: PS00300; SRP54; 1.
 KW Signal recognition particle: GTP-binding; RNA-binding.
 FT DOMAIN 1
 FT DOMAIN 296
 FT NP_BIND 297 534
 FT NP_BIND 109 116
 FT NP_BIND 191 195
 FT NP_BIND 249 252
 FT DOMAIN 468 474
 FT DOMAIN 493 501
 FT DOMAIN 522 533
 FT POLY-GLY.
 FT POLY-GLY.
 SQ SEQUENCE 534 AA; 57117 MW; 1B21E3A48CB3BA6 CRC64;

Query Match 27.0%; Score 50; DB 1; Length 534;
 Best Local Similarity 44.8%; Pred. No. 25;
 Matches 13; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 7 AVEAVAPFGVCYDTKTLGNLNGYAVPNV 35
 DB 236 AFKATADFGAIIITTKTDGHAAGGCAISAV 264

Search completed: May 20, 2003, 16:35:35
 Job time : 12.4231 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 43.2692 seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-13

Perfect score: 185
Sequence: 1 GAPVRAVEAVPAGVCYDRTKLTGNLGGYAVPNV 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	47.0	422	10	08SIV3 oryza sativ
2	79.5	43.0	433	10	09ZVS4 arabidopsis
3	79.5	43.0	434	10	09ZVS5 arabidopsis
4	76	41.1	424	10	08SIV1 oryza sativ
5	74	40.0	419	10	08SIV0 oryza sativ
6	73.5	39.7	452	10	09FSH9 lupinus alb
7	68.5	37.0	413	10	09FEK1 lupinus alb
8	62.5	33.8	449	10	04Z369 lupinus ang
9	59	31.9	433	10	005929 dancus caro
10	58	31.4	402	10	08SIV9 oryza sativ
11	56.5	30.5	525	16	08ZNV0 salmoneilla
12	56.5	30.5	525	16	08ZNV0 salmoneilla
13	56.5	30.5	525	16	08ZNV0 salmoneilla
14	56	30.3	314	10	09FTN5 oryza sativ
15	56	30.3	314	10	08VYH7 oryza sativ
16	55.5	30.0	195	16	09PK21 chlamydia m

17	55.5	30.0	343	7	P79459 ambystoma m
18	55	29.7	369	10	Q9FSZ9 cicer ariet
19	54.5	29.5	594	16	Q9JUN6 neisseria m
20	54	29.2	64	11	Q9JUP3 mus musculu
21	54	29.2	380	10	Q43568 nicotiana t
22	54	29.2	546	10	Q40451 nicotiana t
23	52.5	28.4	238	16	Q8Y2W9 ralsstonia s
24	52.5	28.4	520	16	Q9ZS03 ralsstonia m
25	52	28.1	242	10	Q9SEZ1 streptomyces
26	52	28.1	340	16	Q9X709 streptomyces
27	52	28.1	433	10	Q8RVH5 glycine max
28	52	28.1	541	11	Q60447 cricetus
29	52	28.1	626	16	Q8XQ86 ralsstonia s
30	52	28.1	826	10	Q9SCU3 arabidopsis
31	52	28.1	3198	12	Q9INJ4 pea seed-bo
32	51.5	27.8	478	16	Q9RTM9 deinoecoccus
33	51.5	27.8	520	16	Q97PZ3 streptococc
34	51	27.6	310	5	Q76759 anophelis g
35	51	27.6	327	5	Q76758 anophelis g
36	51	27.6	556	4	Q96S68 homo sapien
37	51	27.6	556	4	Q9BRD6 homo sapien
38	51	27.6	800	11	Q63257 rattus norv
39	51	27.6	1048	10	Q8S249 oryza sativ
40	51	27.6	3199	12	Q8S074 pea seed-bo
41	50.5	27.3	336	16	Q8U6P8 agrobacteri
42	50.5	27.3	495	16	Q34660 bacillus su
43	50.5	27.3	525	16	Q8U1J2 agrobacteri
44	50.5	27.3	594	2	Q51188 neisseria m
45	50.5	27.3	867	3	Q96T27 neurospora

ALIGNMENTS

RESULT 1
ID 08SIV3 PRELIMINARY: PRT: 422 AA.
AC 08SIV3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative dermal glycoprotein.
GN P0504E02.6.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0504E02.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003269; BAB89705.1;
SQ SEQUENCE 422 AA: 44745 MW: 08CF26346400A9F CRC64;

Query Match 47.0%; Score 87; DB 10; Length 422;
Best local Similarity 56.0%; Pred. No. 0.00097;
Matches 14; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

OY 11 VAPFGVCYDRTKLTGNLGGYAVPNV 35
DB 311 VAPFELCYESRTIGSNRLGAYAPDI 335

RESULT 2

ID 09ZVS4 PRELIMINARY: PRT: 433 AA.
AC 09ZVS4;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE F15K9.17 protein (Hypothetical 45.7 kDa protein) (putative
 GN extracellular dermal glycoprotein EDG).
 GN F15K9.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
 RA Kremenetskaia I., Luross J., Araujo R., Buehler E., Conway A.B.,
 RA Dewar K., Feng J., Kim C., Li Y., Shinn P., Sun H., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Theologis A., Ecker J., Davis R.W.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamasu Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shin P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.17 (GI:3850579).";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.17 (GI:3850579).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005278; AAC72119.1; -;
 DR EMBL: AF325092; AAC717160.1; -;
 DR EMBL: AY050988; AAL15204.1; -;
 DR EMBL: AY035026; AKS59531.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 433 AA: 45717 MW: 7214FC4B8BA72962 CRC64;

Query Match 43.0%; Score 79.5; DB 10; Length 433;
 Best Local Similarity 47.1%; Pred. No. 0.011;
 Matches 16; Conservative 6; Mismatches 9; Indels 3; Gaps 1;
 QY 5 ARAVEAVA--PFGVCYDRTKLTGNLGGYAVPNV 35

DB 315 ARSIRKRVASVKGPGACFTKRWGVTRLGYAVPEI 348
 ||::: || ||| : || : || |||| :
 RESULT 3
 ID 09ZVS5 PRELIMINARY; PRT: 434 AA.
 AC 09ZVS5;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE F15K9.16 protein (putative extracellular dermal glycoprotein
 precursor).
 GN F15K9.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
 RA Kremenetskaia I., Luross J., Araujo R., Buehler E., Conway A.B.,
 RA Dewar K., Feng J., Kim C., Li Y., Shinn P., Sun H., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Egu P., Lee J.M.,
 RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
 RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.16 (GI:3850580).";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005278; AAC72120.1; -;
 DR EMBL: AF332411; AAG48774.1; -;
 SQ SEQUENCE 434 AA: 46148 MW: 17DD7684008FDAFC CRC64;

Query Match 43.0%; Score 79.5; DB 10; Length 434;
 Best Local Similarity 47.1%; Pred. No. 0.011;
 Matches 16; Conservative 6; Mismatches 9; Indels 3; Gaps 1;
 QY 5 ARAVEAVA--PFGVCYDRTKLTGNLGGYAVPNV 35
 ||::: || ||| : || : || |||| :
 DB 316 ARSIRKRVASVKGPGACFTKRWGVTRLGYAVPEI 349
 ||::: || ||| : || : || |||| :
 RESULT 4
 ID 08S1V1 PRELIMINARY; PRT: 424 AA.
 AC 08S1V1;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Putative dermal glycoprotein.
 GN P0504E02.8.

OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
 clone: P0504E02.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003269; BAB89707.1;
 SQ SEQUENCE 424 AA; 44620 MW; 0DACD5C8C852C34 CRC64;

Query Match
 Best Local Similarity 41.1%; Score 76; DB 10; Length 424;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 AFGVCYDRTKLTGNNLGYAVPNV 35
 DB 313 APPELCYDSRDLGSTRGLGYAVPQI 336

RESULT 5
 ID Q8S1V0 PRELIMINARY; PRT; 419 AA.
 AC Q8S1V0;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative dermal glycoprotein.
 GN P0504E02.9.
 OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
 clone: P0504E02.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003269; BAB89708.1;
 SQ SEQUENCE 419 AA; 43786 MW; 792BE55A4F3F3A8 CRC64;

Query Match
 Best Local Similarity 40.0%; Score 74; DB 10; Length 419;
 Matches 16; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 8 VEVAPEGVCYDRTKLTGNNLGYAVPNV 35
 DB 307 VAAVAPFLCYRASMLGNTRIGYEPVAV 334

RESULT 6
 ID Q9FSH9 PRELIMINARY; PRT; 452 AA.
 AC Q9FSH9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-NOV-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Conglutin gamma precursor.
 OS *Lupinus albus* (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. ULTRA;

RA Scarafoni A., Di Cataldo A., Vassilevskaia T.D., Bekman E.P.,
 RA Rodrigues-Pousada C., Cecilian F., Duranti M.;
 RT "Cloning, sequencing and expression in seeds and radicles of two
 RT *Lupinus albus* conglutin gamma genes.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ297490; CAC16394.1;
 DR InterPro: IPR001461; AsproteaseA1.
 DR Pfam: PF00026; asp; 1.
 KW Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 452 CONGLUTIN GAMMA.
 SQ SEQUENCE 452 AA; 49219 MW; C343BA283BE752FF CRC64;

Query Match
 Best Local Similarity 39.7%; Score 73.5; DB 10; Length 452;
 Matches 17; Conservative 4; Mismatches 7; Indels 5; Gaps 2;

QY 3 PVARAVEAVAPFGVCYDRTKLTGNNLGYAVPNV 35
 DB 339 PKQAVKAVGPFGLCYDTRKIS--GG--VPSV 366

RESULT 7
 ID Q9FEX1 PRELIMINARY; PRT; 413 AA.
 AC Q9FEX1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Conglutin gamma (Fragment).
 OS *Lupinus albus* (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. ULTRA;
 RA Scarafoni A., Di Cataldo A., Vassilevskaia T.D., Bekman E.P.,
 RA Rodrigues-Pousada C., Cecilian F., Duranti M.;
 RT "Cloning, sequencing and expression in seeds and radicles of two
 RT *Lupinus albus* conglutin gamma genes.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ297568; CAC17729.1;
 FT NON_TER 1 1
 SQ SEQUENCE 413 AA; 45732 MW; 022723A38B16E4B7 CRC64;

Query Match
 Best Local Similarity 37.0%; Score 68.5; DB 10; Length 413;
 Matches 15; Conservative 6; Mismatches 7; Indels 5; Gaps 2;

QY 3 PVARAVEAVAPFGVCYDRTKLTGNNLGYAVPNV 35
 DB 300 PKQAVESVGFGLCYDSRKL--GG--IPSV 327

RESULT 8
 ID Q42369 PRELIMINARY; PRT; 449 AA.
 AC Q42369;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Conglutin gamma precursor.
 OS *Lupinus angustifolius* (Narrow-leaved blue lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3871;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VAR. UNICROP; TISSUE=COFYLEDON;
 RA Gayler K.R.;

```

RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC STRAIN-VAR. UNICROP; TISSUE=COFFLEDON;
RX MEDLINE=93144715; PubMed=8425065;
RA Koliyas S., Gayler K.R.;
RT "Structure of the cDNA coding for conglutin gamma, a sulphur-rich
RL plant Mol. Biol. 21:397-401(1993).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-VAR. UNICROP; TISSUE=COFFLEDON;
RX Gayler K.R.;
RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=95388773; PubMed=7659758;
FA Millar A.A., Kunst L.;
FL Plant Physiol. 108:1747-1747(1995).
CR EMBL; X65601; CAA46552.1; -.
DR EMBL; L39786; AAB53771.1; -.
DR InterPro: IPR001461; AsproteaseA1.
DR Pfam: PF00026; asp_1.
KA Signal.
FT SIGNAL.
FT CHAIN 33 449 POTENTIAL. CONGLUTIN GAMMA.
SQ SEQUENCE 449 AA; 48916 MW; 348BE0FCFBD084F3 CRC64;

Query Match
Best Local Similarity 42.4%; Score 62.5; DB 10; Length 449;
Matches 14; Conservative 6; Mismatches 8; Indels 5; Gaps 2;

OY 3 PVAFAEAVAPFGVCYDTKTLGNLGGYAVPNV 35
DB 336 PKOAVKAVGPFGLCYDSRKIS---GC-APSV 363

RESULT 9
OQ5929 PRELIMINARY; PRT; 433 AA.
ID 005929;
AC 005929;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE EDGP precursor (Fragment).
GN EDGPI.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Saton S.;
RA Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Saton S., Fujii T., Chrispeels M.J.;
RT "cDNA cloning of an extracellular dermal glycoprotein of carrot and
RT its expression in response to wounding.";
RL Planta 188:432-438(1992).
DR EMBL; D14550; BAA03413.1; -.
KW Signal.
FT SIGNAL.
FT CHAIN 21 433 POTENTIAL. EDGP.
SQ SEQUENCE 433 AA; 45841 MW; 06EDFC19B7BDC38E CRC64;

Query Match
Best Local Similarity 39.3%; Score 59; DB 10; Length 433;
Matches 11; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 8 VEAAPFGVCYDTKTLGNLGGYAVPNV 35
DB 323 VASVAPFGACFTDNLSTRLGPSVPSI 350

RESULT 10
OQ8109 PRELIMINARY; PRT; 402 AA.
ID 08109;
AC 08109;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Putative dermal glycoprotein.
GN P0504E02.10.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003269; BAB89709.1; -.
SQ SEQUENCE 402 AA; 41793 MW; E8C84A214D6F6720 CRC64;

Query Match
Best Local Similarity 47.8%; Score 58; DB 10; Length 402;
Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 10 AVAPFGVCYDTKTLGNLGGYAV 32
DB 294 SVAPFELCFNSSALGFTRGYAV 316

RESULT 11
OQ82N60 PRELIMINARY; PRT; 525 AA.
ID 082N60;
AC 082N60;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE GMP synthetase (EC 6.3.5.2).
GN GUA OR STM2510.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008813; ARL21404.1; -.
DR InterPro: IPR000991; GATase_1.
DR InterPro: IPR004739; GMPsynthase_N.
DR InterPro: IPR001674; GMP_synth_C.
DR InterPro: IPR004135; trna_Me_trans.
DR Pfam: PF00117; GATase; 1.
DR Pfam: PF00958; GMP_synth_C; 1.
DR Pfam: PF03054; trna_Me_trans; 1.
DR PRINTS; PR00096; GATASE.

```

DR TIGR00884; guaA_Cterm; 1.
 DR TIGR00888; guaA_Nterm; 1.
 DR PROSITE: PS00442; GATASE_TYPE_I; 1.
 KW ligase: Complete proteome.
 SO SEQUENCE 525 AA; 58716 MW; D354406650969744 CRC64;

Query Match 30.5%; Score 56.5; DB 16; Length 525;
 Best Local Similarity 52.2%; Pred. No. 20;
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 9 EAVAP-FGVCYDRTKLTGNNGY 30
 DB 77 EAGVPVFGVCYGMOTMAMQLGSH 99

RESULT 12

082403 PRELIMINARY; PRT; 525 AA.
 AC 082403;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GMP synthase (glutamine-hydrolyzing) (EC 6.3.5.2).
 GN STY2751.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed-11677608;
 RA Parkhill J., Dougan G., James K.D., Holden N.R., Pickard D., Wain J.,
 Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 Krog A., Larsen T.S., Leather A., Hien T.T., Holroyd S., Jagels K.,
 Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL EMBL: AL627275; CAD02712.1; -;
 DR EMBL: AL627275; CAD02712.1; -;
 DR InterPro: IPR000991; GATase_1.
 DR InterPro: IPR004739; GMPsynthase_N.
 DR InterPro: IPR001674; GMP_synth_C.
 DR InterPro: IPR004135; tRNA_Me_trans.
 DR Pfam: PF00117; GATase; 1.
 DR Pfam: PF00958; GMP_synth_C; 1.
 DR Pfam: PF03054; tRNA_Me_trans; 1.
 DR PRINTS: PR00096; GATASE.
 DR TIGR00884; guaA_Cterm; 1.
 DR TIGR00888; guaA_Nterm; 1.
 DR PROSITE: PS00442; GATASE_TYPE_I; 1.
 KW ligase: Complete proteome.
 SO SEQUENCE 525 AA; 58686 MW; 3086B517F41EA9F9 CRC64;

Query Match 30.5%; Score 56.5; DB 16; Length 525;
 Best Local Similarity 52.2%; Pred. No. 20;
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 9 EAVAP-FGVCYDRTKLTGNNGY 30
 DB 77 EAGVPVFGVCYGMOTMAMQLGSH 99

RESULT 13

08XAB5 PRELIMINARY; PRT; 525 AA.
 AC 08XAB5;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GMP synthase (glutamine-hydrolyzing).
 GN GUA OR 23771 OR ECS3369.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 Grobeck E.J., Davis N.W., Lim A., Dialanta E.T., Potomousis K.,
 Apodaca J., Mantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE005480; AAG57618.1; -;
 DR EMBL: AP002561; BAB36792.1; -;
 DR InterPro: IPR000991; GATase_1.
 DR InterPro: IPR004739; GMPsynthase_N.
 DR InterPro: IPR001674; GMP_synth_C.
 DR InterPro: IPR004135; tRNA_Me_trans.
 DR Pfam: PF00117; GATase; 1.
 DR Pfam: PF00958; GMP_synth_C; 1.
 DR Pfam: PF03054; tRNA_Me_trans; 1.
 DR PRINTS: PR00096; GATASE.
 DR TIGR00884; guaA_Cterm; 1.
 DR TIGR00888; guaA_Nterm; 1.
 DR PROSITE: PS00442; GATASE_TYPE_I; 1.
 KW Complete proteome.
 SO SEQUENCE 525 AA; 58665 MW; 33077AA28660171 CRC64;

Query Match 30.5%; Score 56.5; DB 16; Length 525;
 Best Local Similarity 52.2%; Pred. No. 20;
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 9 EAVAP-FGVCYDRTKLTGNNGY 30
 DB 77 EAGVPVFGVCYGMOTMAMQLGSH 99

RESULT 14

09FTN5 PRELIMINARY; PRT; 314 AA.
 AC 09FTN5;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE putative isoflavone reductase homolog IRL.
 GN P0005A05.14.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0005805.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AP002863; BAB16910.1; -;
 DR InterPro: IPR000875; Cecropin.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003866; Isoflav_reduct.
 DR Pfam: PF02716; Isoflavone_redu. 1.
 DR PROSITE: PS00268; CECROPIN; UNKNOWN.1.
 DR PROSITE: PS00888; CNMP_BINDING_1; UNKNOWN.1.
 SQ SEQUENCE 314 AA; 33501 MW; C6E33803F2E2018 CRC64;

Query Match 30.3%; Score 56; DB 10; Length 314;
 Best Local Similarity 47.1%; Pred. No. 14;
 Matches 16; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

OY 2 APVRAVEAVPFGVCYDTKLGNNLGGYAVPNV 35
 141 AAVRAVEAA--GVY-TYVSNYFAGYALPTI 170

RESULT 15

C8VYH7 PRELIMINARY; PRT; 314 AA.
 ID Q8VYH7;
 AC Q8VYH7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 D3 Isoflavone reductase-like protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RF SEQUENCE FROM N.A.
 RA Kim S.T., Kim S.G., Yu S., Young Y.C., Kang K.Y.;
 RT "Isolation and Characterization of a Rice Isoflavone Reductase (OSIRF)
 RT Gene Expressed by Rice Blast Fungal Elicitor.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY071920; AAL61542.1; -;
 DR InterPro: IPR000875; Cecropin.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003866; Isoflav_reduct.
 DR Pfam: PF02716; Isoflavone_redu. 1.
 DR PROSITE: PS00268; CECROPIN; UNKNOWN.1.
 DR PROSITE: PS00888; CNMP_BINDING_1; UNKNOWN.1.
 SQ SEQUENCE 314 AA; 33487 MW; 534D958039ZCB6C1 CRC64;

Query Match 30.3%; Score 56; DB 10; Length 314;
 Best Local Similarity 47.1%; Pred. No. 14;
 Matches 16; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

OY 2 APVRAVEAVPFGVCYDTKLGNNLGGYAVPNV 35
 141 AAVRAVEAA--GVY-TYVSNYFAGYALPTI 170

Search completed: May 20, 2003, 16:43:07
 Job time : 46.2692 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 16.1538 seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-13
Perfect score: 185
Sequence: 1 GAPVARAVEAVAPFCVCTKTIGNLTGNGYAVPNV 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	29.2	546	2	US-08-669-721-3
2	54	29.2	546	4	US-09-189-344-3
3	51.5	27.8	516	4	US-09-197-063-2
4	50	27.0	534	2	US-08-317-401E-2
5	50	27.0	552	2	US-08-317-401E-4
6	49	26.5	635	4	US-09-341-833A-6
7	49	26.5	635	4	US-09-341-833A-7
8	49	26.5	635	4	US-09-341-833A-9
9	49	26.5	677	4	US-09-341-833A-8
10	46.5	25.1	783	1	US-08-101-593-6
11	46.5	25.1	783	1	US-08-465-995A-6
12	46.5	25.1	783	2	US-08-465-994C-6
13	46	24.9	332	4	US-09-105-390-54
14	46	24.9	1091	3	US-08-633-768A-2
15	46	24.9	1940	4	US-08-644-271-30
16	46	24.9	1940	4	US-09-077-955-34
17	45.5	24.6	393	4	US-09-629-616-2
18	45.5	24.6	523	4	US-09-550-338-2
19	45	24.3	626	4	US-08-961-083-220
20	45	24.3	1121	1	US-07-789-915A-2
21	45	24.3	1121	1	US-08-005-002C-2
22	45	24.3	1121	1	US-08-487-203A-2
23	45	24.3	1334	6	5476657-1
24	45	24.3	1403	1	US-07-908-253-3
25	45	24.3	1403	2	US-08-694-865-17
26	45	24.3	1403	2	US-08-535-837-3
27	45	24.3	1403	3	US-09-124-491-17

28	44	23.8	332	4	US-09-105-390-58	Sequence 58, Appl
29	44	23.8	399	3	US-08-651-999A-3	Sequence 3, Appl
30	44	23.8	399	4	US-09-385-752-3	Sequence 3, Appl
31	44	23.8	731	2	US-08-911-364-1	Sequence 1, Appl
32	44	23.8	733	4	US-08-464-700-2	Sequence 2, Appl
33	44	23.8	765	4	US-09-134-001C-3681	Sequence 3681, Ap
34	44	23.8	792	2	US-08-678-039A-40	Sequence 40, Appl
35	44	23.8	15281	2	US-08-471-119A-2	Sequence 2, Appl
36	43.5	23.5	332	4	US-09-064-411A-39	Sequence 39, Appl
37	43.5	23.5	366	4	US-09-064-411A-37	Sequence 37, Appl
38	43	23.2	212	2	US-08-663-566A-7	Sequence 7, Appl
39	43	23.2	212	2	US-08-023-610-7	Sequence 7, Appl
40	43	23.2	212	2	US-08-288-065A-7	Sequence 7, Appl
41	43	23.2	212	2	US-08-362-240A-7	Sequence 7, Appl
42	43	23.2	212	4	US-08-804-372A-5	Sequence 5, Appl
43	43	23.2	212	5	PCT-US95-10245-7	Sequence 7, Appl
44	43	23.2	331	4	US-09-105-390-56	Sequence 56, Appl
45	43	23.2	370	1	US-07-966-187-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-669-721-3
: Sequence 3, Application US/08669721
: Patent No. 5834236
:
: GENERAL INFORMATION:
: APPLICANT: Lamb et al., Christopher J.
: TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
: TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/669,721
: FILING DATE: 27-JUN-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ellison, Eldora L.
: REGISTRATION NUMBER: 39,967
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 546 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-669-721-3
:
: Query Match 29.2% Score 54; DB 2; Length 546;
: Best Local Similarity 35.3%; Pred. No. 58;
: Matches 12; Conservative 3; Mismatches 19; Indels 0; Gaps 0;
```

QY 1 GAPVARAVEAVAPFCVCTKTIGNLTGNGYAVPN 34
DB 400 GPPATVGTVDVPIAARDTENTLPAVGGCGVTN 433

RESULT 2

US-09-189-344-3
; Sequence 3, Application US/09189344
; Patent No. 6191258
; GENERAL INFORMATION:
; APPLICANT: Lamb et al., Christopher J.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
; NUMBER OF INVENTIONS: 22
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/189,344
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/669,721
; FILING DATE: 27-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellison, Eldora L.
; REGISTRATION NUMBER: 39,967
; REFERENCE/DOCKET NUMBER: 07251/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-189-344-3

Query Match 29.2%; Score 54; DB 4; Length 546;
Best Local Similarity 35.3%; Pred. No. 5.8;
Matches 12; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

QY 1 GAVRAAEVAVPFGVCYDTKLTGNNLGGYAVPN 34
DB 400 GPPATVGVTDVPIAAFTENLPPNAVGGGVYN 433

RESULT 3
US-09-197-063-2
; Sequence 2, Application US/09197063
; Patent No. 6261817
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; APPLICANT: Biswas, Sanjoy
; APPLICANT: Warren, Richard L.
; APPLICANT: Shilling, Lisa K.
; TITLE OF INVENTION: NO. 6261817el Guaa
; FILE REFERENCE: GM10121
; CURRENT APPLICATION NUMBER: US/09/197,063
; CURRENT FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/066,350
; NUMBER OF SEQ ID NOS: 5
; SCFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

US-09-197-063-2
Query Match 27.8%; Score 51.5; DB 4; Length 516;
Best Local Similarity 33.3%; Pred. No. 13;
Matches 11; Conservative 4; Mismatches 9; Indels 9; Gaps 1;

QY 1 GAVRAAEVAVPFGVCYDTKLTGNNLGGYAVPN 33
DB 79 GPI-----LGIYGMQLTTHKLGKVP 102

RESULT 4
US-08-317-401E-2
; Sequence 2, Application US/08317401E
; Patent No. 5922561
; GENERAL INFORMATION:
; APPLICANT: Thompson, Sheryl Ann
; APPLICANT: Yaver, Debbie Sue
; TITLE OF INVENTION: GENES ENCODING SIGNAL RECOGNITION PARTICLE OF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5922561o No. 5922561disk of No. 5922561th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,401E
; FILING DATE: 03-October-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 4248.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; US-08-317-401E-2

Query Match 27.0%; Score 50; DB 2; Length 534;
Best Local Similarity 44.8%; Pred. No. 23;
Matches 13; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 7 AVEAVAPFGVCYDTKLTGNNLGGYAVPNV 35
DB 236 ARKATADFGAIIITTKDGHAGGAIASV 264

RESULT 5
US-08-317-401E-4
; Sequence 4, Application US/08317401E
; Patent No. 5922561
; GENERAL INFORMATION:
; APPLICANT: Thompson, Sheryl Ann
; APPLICANT: Yaver, Debbie Sue
; TITLE OF INVENTION: GENES ENCODING SIGNAL RECOGNITION PARTICLE OF
; NUMBER OF SEQUENCES: 10

Sequence 8, Application US/09341833A
Patent No. 6383488
GENERAL INFORMATION:
APPLICANT: CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA
TITLE OF INVENTION: PRE-M/M PROTEIN EPTIOPES OF THE DENGUE VIRUS, SYNTHETIC
FILE REFERENCE: Dengue prem/M SEOS 1-9 for 976-4(0003)
CURRENT APPLICATION NUMBER: US/09/341,833A
PRIOR FILING DATE: 1999-10-14
PRIOR APPLICATION NUMBER: CU 13/97
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 677
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
OTHER INFORMATION: Of Dengue virus prem/M and Neisseria meningitidis
US-09-341-833A-8

Query Match 26.5%; Score 49; DB 4; Length 677;
Best Local Similarity 37.5%; Pred. No. 43;
Matches 15; Conservative 2; Mismatches 13; Indels 10; Gaps 2;

QY 2 GAPVARAVEAVAP-----FG-----VCYDFTKLTGNNLGGYVA 31
DB 175 GAPAEAPKAAAPAPQAQFGSGADAEYDVYVLCGGPGGYS 214

RESULT 10
US-08-101-593-6
Sequence 6, Application US/08101593
Patent No. 5547859
GENERAL INFORMATION:
APPLICANT: Goodman, Myron F.
APPLICANT: Reha-Krantz, Linda J.
TITLE OF INVENTION: NEW DNA SEQUENCING ENZYMES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson
STREET: 201 No. 5547859th Figueroa Street, Fifth Floor
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/101,593
FILING DATE: 19930802
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Spitals, John P.
REGISTRATION NUMBER: 29,215
REFERENCE/DOCKET NUMBER: 1920-305
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 977-1001
TELEFAX: (213) 977-1003
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-101-593-6

Query Match 25.1%; Score 46.5; DB 1; Length 783;

Best Local Similarity 37.1%; Pred. No. 1.2e+02;
Matches 13; Conservative 4; Mismatches 15; Indels 3; Gaps 1;
QY 1 GAPVARAVEAVAPFGCYDFTKLTGNNLGGYVAPNV 35
DB 360 GLPVDHRGSAVAFGLHYPPRM--HRAGVAPNL 391

RESULT 11
US-08-465-995A-6
Sequence 6, Application US/08465995A
Patent No. 5660980
GENERAL INFORMATION:
APPLICANT: Myron F. Goodman
APPLICANT: Linda J. Reha-Krantz
TITLE OF INVENTION: METHODS FOR IDENTIFYING AND ISOLATING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson, LLP
STREET: 201 No. 5660980th Figueroa Street, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: ASCII DOS/TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,995A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Robert Berliner
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 1920-30502
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/977-1001
TELEFAX: 213/977-1003
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-995A-6

Query Match 25.1%; Score 46.5; DB 1; Length 783;
Best Local Similarity 37.1%; Pred. No. 1.2e+02;
Matches 13; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

QY 1 GAPVARAVEAVAPFGCYDFTKLTGNNLGGYVAPNV 35
DB 360 GLPVDHRGSAVAFGLHYPPRM--HRAGVAPNL 391

RESULT 12
US-08-465-994C-6
Sequence 6, Application US/08465994C
Patent No. 5928919
GENERAL INFORMATION:
APPLICANT: MYRON F. GOODMAN
APPLICANT: LINDA L. REHA-KRANTZ
TITLE OF INVENTION: VARIANT DNA POLYMERASES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson, LLP
STREET: 201 No. 5928919th Figueroa Street, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: ASCII DOS/TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,994C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MARGARET A. CHURCHILL
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1920-305D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/977-1001
TELEFAX: 213/977-1003
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-994C-6

Query Match 25.1% Score 46.5; DB 2; Length 783;
Best Local Similarity 37.1%; Pred. No. 1.2e+02;
Matches 13; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

QY 1 GAPVARAVAPFGVCYDTKLTGNLGGYAVPNV 35
DB 360 GLPVDHGGSAVAFGLHYPRM--HRAQYVAPNL 391

RESULT 13
US-09-105-390-54
Sequence 54, Application US/09105390
Patent No. 6288303
GENERAL INFORMATION:
APPLICANT: RODRIGUEZ, Raymond
TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
TITLE OF INVENTION: and Genes
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,390
FILING DATE: Filed herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,675
FILING DATE: 25-JUN-97
ATTORNEY/AGENT INFORMATION:
NAME: Petlthory, Joanne R.
REGISTRATION NUMBER: P42,995
REFERENCE/DOCKET NUMBER: 2000-0455.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-105-390-54

Query Match 24.9% Score 46; DB 4; Length 332;
Best Local Similarity 52.2%; Pred. No. 50;
Matches 12; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 5 ARRAVAVAPFGVCYDTKLTGNL 27
DB 21 ASAPTVOISIGVCYG--VLGNL 41

RESULT 14
US-08-633-768A-2
Sequence 2, Application US/08633768A
Patent No. 6013504

GENERAL INFORMATION:
APPLICANT: YU, SHUKUN
APPLICANT: BOISEN, KIRSTEN
APPLICANT: KRAGH, KARSTEN
APPLICANT: BOJKO, MAJA
APPLICANT: NIELSEN, JOHN
APPLICANT: MARCUSSEN, JAN
TITLE OF INVENTION: ALPHA-1,4-GLUCAN LYASE FROM
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,768A
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321301.5
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DY007,001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1091 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-768A-2

Query Match 24.9% Score 46; DB 3; Length 1091;
Best Local Similarity 46.2%; Pred. No. 2.2e+02;
Matches 12; Conservative 2; Mismatches 6; Indels 6; Gaps 1;

QY 14 FGVCYDTKLTGNLGG-----YAVP 33
DB 896 KCPDYDKRLDSALDGGQMIKNYSVP 921

```

RESULT 15
US-08-644-271-30
: Sequence 30 Application US/08644271
: Patent No. 5814478
: GENERAL INFORMATION:
: APPLICANT: Valenzuela, et al.
: TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
: TITLE OF INVENTION: AND LIGANDS
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Regeneron Pharmaceuticals, Inc.
: STREET: 777 Old Saw Mill Road
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/644,271
: FILING DATE: 10-MAY-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: USSN 60/008,657
: FILING DATE: 15-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Covert, Robert J
: REGISTRATION NUMBER: 36,108
: REFERENCE/DOCKET NUMBER: REG 195A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 914-345-7400
: TELEFAX: 914-345-7721
: TELEX:
: INFORMATION FOR SEQ ID NO: 30:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1940 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Rat Agrin
: LOCATION: 1..1940
: OTHER INFORMATION:
: US-08-644-271-30

Query Match 24.9%; Score 46; DB 2; Length 1940;
Best Local Similarity 45.5%; Pred.No. 4.4e+02;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0.

OY 1 GAPVARAVEAVAPGVCYDTKT 22
|||:||||:|||||
db 990 GPTIERASCYNSPLGCCSDGKT 1011

```

Search completed: May 20, 2003, 16:44:37
Job time : 18.1538 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 ; Search time 17.3077 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-13
Perfect score: 185
Sequence: 1 GAPVAAVEAAVAFGVCYDFTKTLGNLGGVAVPNV 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCIT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56.5	30.5	525	10	US-09-815-242-10249
2	56.5	30.5	525	10	US-09-815-242-11741
3	56.5	30.5	525	10	US-09-815-242-13950
4	54	29.2	546	10	US-09-839-743-3
5	52.5	28.4	525	10	US-09-815-242-11952
6	52	28.1	142	9	US-09-738-626-6400
7	51.5	27.8	520	10	US-09-815-242-13283
8	49.5	26.8	520	10	US-09-815-242-10605
9	48	25.9	267	9	US-09-813-408-22
10	48	25.9	712	10	US-09-815-242-10930
11	47.5	25.7	523	10	US-09-815-242-10987
12	46.5	25.1	719	9	US-10-270-875-34
13	46.5	25.1	719	9	US-10-270-878-34
14	46.5	25.1	719	9	US-10-270-786-34
15	46.5	25.1	719	9	US-10-270-710-34
16	46.5	25.1	719	9	US-10-270-859-34
17	46.5	25.1	1092	9	US-09-423-126-5
18	46	25.9	266	10	US-09-837-235-18
19	46	24.9	588	9	US-09-908-193-34

20	46	24.9	695	10	US-09-764-898-191	Sequence 191, App
21	46	24.9	778	9	US-09-908-193-8	Sequence 8, Appl
22	46	24.9	779	9	US-09-908-193-10	Sequence 10, Appl
23	46	24.9	1091	9	US-09-423-126-2	Sequence 2, Appl
24	46	24.9	1091	10	US-09-280-197-2	Sequence 2, Appl
25	46	24.9	1940	9	US-10-016-283-34	Sequence 34, Appl
26	45.5	24.6	177	9	US-09-791-279-148	Sequence 148, App
27	45.5	24.6	393	9	US-09-738-626-5280	Sequence 5280, Ap
28	45.5	24.6	393	9	US-10-284-138-2	Sequence 2, Appl
29	45.5	24.6	393	9	US-10-284-334-2	Sequence 2, Appl
30	45.5	24.6	393	10	US-09-836-4708-2	Sequence 2, Appl
31	45.5	24.6	523	9	US-09-738-626-4177	Sequence 4177, Ap
32	45.5	24.6	523	10	US-09-767-878-2	Sequence 2, Appl
33	45	24.3	577	10	US-09-758-269-10	Sequence 220, App
34	45	24.3	626	10	US-09-755-272-220	Sequence 2, Appl
35	45	24.3	1010	10	US-09-759-152-2	Sequence 2, Appl
36	45	24.3	1024	10	US-09-879-257A-30	Sequence 30, Appl
37	45	24.3	5179	9	US-10-025-380-1068	Sequence 1068, Ap
38	45	24.3	5179	10	US-09-922-217-1068	Sequence 1068, Ap
39	45	24.3	5179	10	US-09-833-263-1068	Sequence 830, App
40	44.5	24.1	103	9	US-09-796-692-830	Sequence 2420, Ap
41	44.5	24.1	103	9	US-09-796-692-2420	Sequence 830, App
42	44.5	24.1	103	9	US-10-040-862-830	Sequence 830, App
43	44.5	24.1	103	9	US-10-040-862-2420	Sequence 2420, Ap
44	44.5	24.1	203	10	US-09-809-545A-2	Sequence 2, Appl
45	44.5	24.1	372	10	US-09-873-880-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-10249
Sequence 10249, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA-011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTED for Windows Version 4.0
SEQ ID NO 10249
LENGTH: 525
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10249
Query Match 30.5%; Score 56.5; DB 10; Length 525;
Best Local Similarity 52.2%; Pred. No. 3.4;

Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 9 EAVAP-FGVCYDTKTLGNLCGY 30
||| ||||| :|: |||:
Db 77 EAGVPFVGVCYGMOTMAMOLGCH 99

RESULT 2
US-09-815-242-11741
Sequence 11741, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11741
LENGTH: 525
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(525)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-11741

Query Match 30.5%; Score 56.5; DB 10; Length 525;
Best Local Similarity 52.2%; Pred. No. 3.4;
Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 9 EAVAP-FGVCYDTKTLGNLCGY 30
||| ||||| :|: |||:
Db 77 EAGVPFVGVCYGMOTMAMOLGCH 99

RESULT 3
US-09-815-242-13950
Sequence 13950, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13950
LENGTH: 525
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13950

Query Match 30.5%; Score 56.5; DB 10; Length 525;
Best Local Similarity 52.2%; Pred. No. 3.4;
Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 9 EAVAP-FGVCYDTKTLGNLCGY 30
||| ||||| :|: |||:
Db 77 EAGVPFVGVCYGMOTMAMOLGCH 99

RESULT 4
US-09-839-743-3
Sequence 3, Application US/09839743
Patent No. US20020146824A1
GENERAL INFORMATION:
APPLICANT: The Salk Institute for Biological Sciences
APPLICANT: Lamb, Christopher
APPLICANT: Doerner, Peter
APPLICANT: Labile, Goetz
TITLE OF INVENTION: NO. US20020146824A1el Transcription Enhancer Element and
FILE REFERENCE: SALKINS.008DV3
CURRENT APPLICATION NUMBER: US/09/839,743
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/401,336
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: US 09/189,344
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/669,721
PRIOR FILING DATE: 1996-06-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 546
TYPE: PRT
ORGANISM: Nicotiana tabacum
US-09-839-743-3

Query Match 29.2%; Score 54; DB 10; Length 546;
Best Local Similarity 35.3%; Pred. No. 8.5;
Matches 12; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GAVPAAVEAVAPFGVCYDTKTLGNLCGYAVPN 34
||| ||||| :|: |||:
Db 400 GPPAAIVGVTDVPIAAAFDTEMLPNAYGGCVTN 433

```

1      RESULT 5
2      US-09-815-242-11952
3      ; Sequence 11952, Application US/09815242
4      ; Patent No. US20020061569A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Haselbeck, Robert
7      ; APPLICANT: Ohlsen, Kari L.
8      ; APPLICANT: Zyskind, Judith W.
9      ; APPLICANT: Wall, Daniel
10     ; APPLICANT: Trawick, John D.
11     ; APPLICANT: Carr, Grant J.
12     ; APPLICANT: Yamamoto, Robert T.
13     ; TITLE OF INVENTION: Identification of Essential Genes In
14     ; TITLE OF INVENTION: Prokaryotes
15     ; FILE REFERENCE: ELITRA.011A
16     ; CURRENT FILING DATE: 2001-03-21
17     ; PRIOR APPLICATION NUMBER: 60/191,078
18     ; PRIOR FILING DATE: 2000-03-21
19     ; PRIOR APPLICATION NUMBER: 60/206,848
20     ; PRIOR FILING DATE: 2000-05-23
21     ; PRIOR APPLICATION NUMBER: 60/207,727
22     ; PRIOR FILING DATE: 2000-05-26
23     ; PRIOR APPLICATION NUMBER: 60/242,578
24     ; PRIOR FILING DATE: 2000-10-23
25     ; PRIOR APPLICATION NUMBER: 60/253,625
26     ; PRIOR FILING DATE: 2000-11-27
27     ; PRIOR APPLICATION NUMBER: 60/257,931
28     ; PRIOR FILING DATE: 2000-12-22
29     ; PRIOR APPLICATION NUMBER: 60/269,308
30     ; PRIOR FILING DATE: 2001-02-16
31     ; NUMBER OF SEQ ID NOS: 14110
32     ; SOFTWARE: FastSeq for Windows Version 4.0
33     ; SEQ ID NO 11952
34     ; LENGTH: 525
35     ; TYPE: PRT
36     ; ORGANISM: Pseudomonas aeruginosa
37
38      US-09-815-242-11952
39
40      Query Match      28.4%; Score 52.5; DB 10; Length 525;
41      Best Local Similarity 41.4%; Pred. No. 14;
42      Matches 12; Conservative 5; Mismatches 7; Indels 5; Gaps 2
43
44      QY      6 RAVEAV---AP-FGVCTDTKTLGNLGG 29
45              11:11      11:11      11:      :11
46
47      Db      70 RAPQAVFDLKVPLFGICVGMQTMARQMG 98
48
49      RESULT 6
50      US-09-738-626-6400
51      ; Sequence 6400, Application US/09738626
52      ; Publication No. US20020197605A1
53      ; GENERAL INFORMATION:
54      ; APPLICANT: NAKAGAWA, SATOSHI
55      ; APPLICANT: MIZOGUCHI, HIROSHI
56      ; APPLICANT: ANDO, SEIKO
57      ; APPLICANT: HAYASHI, MIKIRO
58      ; APPLICANT: OCHIAI, KEIKO
59      ; APPLICANT: YOKOI, HARUHIKO
60      ; APPLICANT: TATEISHI, NAKO
61      ; APPLICANT: SENOH, AKIHIRO
62      ; APPLICANT: IKEDA, MASATO
63      ; APPLICANT: OZAKI, AKIO
64      ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
65      ; FILE REFERENCE: 249-125
66      ; CURRENT APPLICATION NUMBER: US/09/738,626
67      ; CURRENT FILING DATE: 2000-12-18
68      ; PRIOR APPLICATION NUMBER: JP 99/377484
69      ; PRIOR FILING DATE: 1999-12-16
70      ; PRIOR APPLICATION NUMBER: JP 00/159162
71      ; PRIOR FILING DATE: 2000-04-07
72      ; PRIOR APPLICATION NUMBER: JP 00/280988

```

```

PRIORITY FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6400
LENGTH: 142
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6400

Query Match      28.1%; Score 52; DB 9; Length 142;
Best Local Similarity 33.3%; Pred. No. 3.6;
Matches 10; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Oy      1 GAPVAAVEAVALPFCVCDTKTLGNNGCY 30
      | | : : : | | | : : : | : | :
Db      65 GDPVATIGLASFGVAIARIPGLVMOH 94

RESULT 7
US-09-815-242-13283
Sequence 13283, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13283
LENGTH: 520
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13283

Query Match      27.8%; Score 51.5; DB 10; Length 520;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 11; Conservative 4; Mismatches 9; Indels 9; Gaps 1;

Oy      1 GAPVAAVEAVALPFCVCDTKTLGNNGCYAMP 33
      | | : : : | | | : : : | : | :
Db      83 GPII-----LGICYGMDLTHRLGKVP 106

RESULT 8
US-09-815-242-10605
Sequence 10605, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert

```

```
US-09-815-242-10930
: Sequence 10930, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIORITY FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10930
: LENGTH: 712
: TYPE: PRF
: ORGANISM: Enterococcus faecalis
US-09-815-242-10930

Query Match      25.9%; Score 48; DB 10; Length 712;
Best Local Similarity 43.5%; Pred. No. 90;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0

QY      I GAPARAWEAVAPFCVCYDFTKL 23
       | :||| | : : : : ||
Db      170 GTLYAKAVEKVPETIAFDTRL 192

RESULT 11
US-09-815-242-10987
: Sequence 10987, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIORITY FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10930
: LENGTH: 712
: TYPE: PRF
: ORGANISM: Enterococcus faecalis
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 34
; LENGTH: 719
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-270-710-34

Query Match 25.1%; Score 46.5; DB 9; Length 719;
Best Local Similarity 37.1%; Pred. NO. 1.5e+02;
Matches 13; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

QY 1 GAPVAVAEVAPFGVCYCDTKTLGNLGGYAVPNV 35
DB 308 GLEVDRHGGVAAFGHLYFPRM--HRAGVAVPNL 339

Search completed: May 20, 2003, 18:00:00
Job time: 18.3077 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 ; Search time 18.4945 Seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-14

Perfect score: 96

Sequence: 1 KRLGSRHPHTGCGGL 17

Scoring table:

BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	17	AAV93757	Amino acid sequenc
2	96	100.0	17	AAE14688	Wheat flour xylana
3	96	100.0	17	AAU07393	Bacillus subtilis
4	96	100.0	196	AAU75817	Wheat L endoxylana
5	96	100.0	370	AAU75815	Wheat L endoxylana
6	96	100.0	370	AAU75816	Wheat L endoxylana
7	96	100.0	381	AAU75812	Wheat L endoxylana
8	96	100.0	381	AAU75813	Wheat L endoxylana
9	96	100.0	381	AAU75814	Wheat L endoxylana
10	61	63.5	14	AAU75829	Wheat xylanase inh

11	46.5	48.4	198	21	AA624723	Arabidopsis thalia
12	46.5	48.4	198	21	AA645304	Arabidopsis thalia
13	46.5	48.4	252	21	AA624722	Arabidopsis thalia
14	46.5	48.4	252	21	AA645303	Arabidopsis thalia
15	46.5	48.4	256	21	AA624721	Arabidopsis thalia
16	46.5	48.4	256	21	AA645302	Arabidopsis thalia
17	45	46.9	233	22	AA661608	Human protein HP03
18	45	46.9	276	23	AAE14745	Human triacylglyce
19	45	46.9	280	23	AAE14743	Human triacylglyce
20	45	46.9	289	23	AAU77494	Human lipid metabo
21	45	46.9	395	23	AAE14744	Human triacylglyce
22	45	46.9	397	23	AAE17308	Human lysosomal ac
23	45	46.9	399	22	AA667513	Amino acid sequenc
24	45	46.9	399	23	AAU77496	Human lipid metabo
25	45	46.9	473	22	AAU40252	Protonibacterium
26	45	46.9	656	22	ABG26839	Novel human diarno
27	45	46.9	666	21	AA642538	Human ORF2302
28	45	46.9	731	23	AAE14746	Human triacylglyce
29	44.5	46.4	218	22	AA690314	C glutamincum prote
30	44	45.8	477	22	ABG12357	Novel human diarno
31	44	45.8	932	19	AAW53281	Murine phospholipa
32	43.5	45.3	158	21	AAV58264	Fennugreek galactom
33	43.5	45.3	438	21	AAV58262	Fennugreek galactom
34	43.5	45.3	472	21	AAV58266	Fennugreek galactom
35	43	44.8	163	23	ABB04853	LDL receptor bindi
36	42	43.8	159	22	AAW04665	Human polypeptide
37	42	43.8	163	23	ABB04848	LDL receptor bindi
38	42	43.8	325	21	AA623196	Arabidopsis thalia
39	42	43.8	325	21	AA632391	Arabidopsis thalia
40	42	43.8	463	21	AA607155	Arabidopsis thalia
41	42	43.8	463	21	AA648101	Arabidopsis thalia
42	42	43.8	494	21	AA648093	Arabidopsis thalia
43	42	43.8	844	22	ABG08453	Novel human diarno
44	41	42.7	41	21	AAV65359	Human 5' EST relat
45	41	42.7	58	22	AAU54989	Protonibacterium

ALIGNMENTS

RESULT 1
AAV93757
ID AAV93757 standard; peptide: 17 AA.
XX
AC AAV93757;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.
XX
KW Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW dough; dough preparation.
XX
OS Triticum sp.
XX
PN WO200039289-A2.
XX
PD 06-JUL-2000.
XX
PF 17-DEC-1999; 99WO-1B02071.
XX
PR 23-DEC-1998; 98GB-0028599.
PR 06-APR-1999; 99GB-0007805.
PR 15-APR-1999; 99GB-0008645.
XX
PA (DANI-) DANISCO AS.
XX
PI Sibbesen O, Sorensen JF;
XX
DR WPI; 2000-465744/40.
XX
PT Mutant xylanase protein identified using xylanase inhibitor useful for
preparing non-sticky dough for bakery products -

XX Claim 24; Page 112; 112pp; English.
 PS
 CC The present sequence is derived from an endo-beta-1,4-xylanase
 CC inhibitor. The protein is obtained from wheat flour. The specification
 CC also describes a mutant xylanase protein. The xylanase is useful for
 CC preparing a foodstuff, preferably a bakery product or a substance
 CC (e.g. a dough) for making the bakery product. Wild type xylanase or
 CC mutant xylanase is useful for preparing a dough that is less sticky
 CC than a dough comprising a fungal xylanase. The xylanase inhibitor is
 CC useful for screening high degree resistance xylanases for dough
 CC preparation. The xylanase is also useful for preparing a non-sticky
 CC dough. A combination of xylanase and the inhibitor is useful for
 CC calibrating and/or determining the quantity of inhibitor in a wheat
 CC flour sample.
 CC
 SO Sequence 17 AA;
 XX
 XX
 Query Match 100.0%; Score 96; DB 21; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KRIGFSRLPHFTGCGGL 17
 DB 1 KRIGFSRLPHFTGCGGL 17
 RESULT 2
 AAE14688
 ID AAE14688 standard; peptide; 17 AA.
 AC AAE14688;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE wheat flour xylanase inhibitor A chain C-terminal fragment.
 XX
 KW Refrigerated dough; syruping; arabinoxylan; bakery product; bread;
 KW pizza base; cake; biscuit; wheat; flour; xylanase inhibitor.
 OS Trifolium aestivum.
 XX
 PN WO200152657-A1.
 PD 26-JUL-2001.
 XX
 PF 17-JAN-2001; 2001WO-1B00168.
 XX
 PR 18-JAN-2000; 2000GB-0001136.
 PA (DANI-) DANISCO AS.
 PS
 PI Poulsen CH, Sorensen JF;
 XX
 DR WPI; 2001-457446/49.
 PT Production of refrigerated dough with reduced syruping, useful in
 PT production of bakery products such as bread, comprises admixing cereal
 PT flour, water and protein that prevents enzymatic degradation of
 PT arabinoxylan in the cereal flour -
 XX
 PS Disclosure; Page 23; 26pp; English.
 XX
 CC The invention relates to a process for producing refrigerated dough
 CC with reduced 'syruping' (precipitation of liquid on the dough surface
 CC because of a reduction in water holding capacity caused by the breakdown
 CC of arabinoxylan over time). The process comprises admixing cereal flour
 CC and water with a protein that reduces/prevents enzymatic degradation of
 CC arabinoxylan in the cereal flour. The preferred protein is a xylanase
 CC inhibitor. The method is useful to produce refrigerated dough in which
 CC syruping is reduced or eliminated. Refrigerated dough is typically
 CC stored for long periods to enable fresh baked products (e.g. bread,
 CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
 CC requirements using the method by the use of specific proteins/protein
 CC combinations. The present sequence is wheat flour
 CC endo-beta-1,4-xylanase inhibitor A chain C-terminal fragment.
 XX
 SO Sequence 17 AA;
 XX
 XX
 *Query Match 100.0%; Score 96; DB 22; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KRIGFSRLPHFTGCGGL 17
 DB 1 KRIGFSRLPHFTGCGGL 17
 RESULT 3
 AAU07393
 ID AAU07393 standard; protein; 17 AA.
 XX
 AC AAU07393;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Bacillus subtilis xylanase inhibitor #2.
 XX
 KW Xylanase; plant cell wall; baking; cereal; starch production; wood;
 KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
 OS Bacillus subtilis.
 XX
 PN WO200166711-A1.
 PD 13-SEP-2001.
 XX
 PF 08-MAR-2001; 2001WO-1B00426.
 XX
 PR 08-MAR-2000; 2000GB-0005585.
 PR 27-JUN-2000; 2000GB-0015751.
 PA (DANI-) DANISCO AS.
 PS
 PI Sibbesen O, Sorensen JF;
 XX
 DR WPI; 2001-596834/67.
 XX
 PT Novel variant xylanase polypeptide or its fragment useful for degrading
 PT or modifying plant cell wall, comprises amino acid modifications such
 PT that the polypeptide has altered sensitivity to xylanase inhibitor -
 XX
 PS Disclosure; Page 63; 70pp; English.
 XX
 CC The invention relates to a variant xylanase polypeptide (I) or its
 CC fragment having xylanase activity, comprising one or more amino acid
 CC modifications such that (I) or its fragment has an altered sensitivity to
 CC a xylanase inhibitor as compared with the parent xylanase enzyme. (I) or
 CC its coding sequence (II) is useful for degrading or modifying plant cell
 CC wall or for processing a plant material by contacting the plant cell wall
 CC or plant material with (I) or (II). (I) is useful for modifying plant
 CC materials, and in baking, processing cereals, starch production,
 CC processing wood and enhancing the bleaching of wood pulp. (II)
 CC is useful for altering the viscosity derived from the presence of
 CC hemicellulose or arabinoxylan in a solution or system comprising plant
 CC cell wall material. (I) is useful for preparing a foodstuff such as
 CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
 CC The present sequence represents the amino acid sequence of Bacillus
 CC subtilis xylanase inhibitor #2 as described in the method of the
 CC invention.
 XX
 SO Sequence 17 AA;
 XX
 XX
 Query Match 100.0%; Score 96; DB 22; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRIGSRPLPFTGCGGL 17
 1 KRIGSRPLPFTGCGGL 17

Db 1 KRIGSRPLPFTGCGGL 17

RESULT 4
 AAU75817
 ID AAU75817 standard: Protein; 196 AA.
 XX AAU75817;
 AC
 XX
 XX 23-APR-2002 (first entry)
 DE Wheat L endoxylanase inhibitor, TDXI I, partial sequence TDXI-I.01.
 XX
 XX Wheat; TDXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Triticum durum cultivar Mexicali.
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 XX Delcour J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR N-PSDB: ABK13674.
 DR WPI: 2002-114579/15.
 XX
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.

The invention relates to separating and/or isolating inhibitors of cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulose, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syruping in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TDXI-I (T. durum L endoxylanase
 CC inhibitor).
 XX
 XX SQ Sequence 196 AA:

Query Match 100.0%; Score 96; DB 23; Length 196;
 Best Local Similarity 100.0%; Pred. No. 1.9e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRIGSRPLPFTGCGGL 17
 1 KRIGSRPLPFTGCGGL 196

Db 180 KRIGSRPLPFTGCGGL 196

RESULT 5
 AAU75815
 ID AAU75815 standard: Protein; 370 AA.
 XX
 XX AAU75815;
 AC
 XX
 XX 23-APR-2002 (first entry)
 DE
 XX Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.
 XX
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 XX Triticum aestivum cultivar Solissons.
 OS
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 XX Delcour J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR N-PSDB: ABK13672.
 DR WPI: 2002-114579/15.
 XX
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.

The invention relates to separating and/or isolating inhibitors of cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising screening the inhibition activity by using two or more enzymes during the

XX	plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
OS	Triticum aestivum.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 95
FT	/Label= Unknown
FT	Misc-difference 98
FT	/Label= Unknown
FT	Misc-difference 101
FT	/Label= Unknown
FT	Misc-difference 110
FT	/Label= Unknown
FT	Misc-difference 333
FT	/Label= Unknown
XX	
PN	WO200198474-A1.
PD	
XX	27-DEC-2001.
XX	
FE	21-JUN-2001; 2001WO-BE00106.
XX	
PR	22-JUN-2000; 2000GB-0015296.
PR	25-JAN-2001; 2001GB-0002018.
PR	26-JAN-2001; 2001GB-0002194.
PR	16-MAR-2001; 2001GB-0005564.
XX	21-MAY-2001; 2001GB-0012328.
PA	(LEUV-) LEUVEN RES & DEV.
PI	
PI	Delcour J, Debayer W, Gebruers K, Goesaert H, Flerens K, Robben J;
DR	Van Campenhout S;
XX	WPI: 2002-114579/15.
PT	
PT	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT	beta-glucanolytic enzymes comprises using endoxylanases during
PT	screening for inhibition activity or affinity chromatography with
PT	immobilised enzymes -
PS	
XX	Claim 127: Page 50-51; 127pp: English.
XX	
CC	The invention relates to separating and/or isolating inhibitors of
CC	cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
CC	screening the inhibition activity by using two or more enzymes during the
CC	separation and/or isolation steps that allow to distinguish inhibitors of
CC	different specificity or by using an affinity chromatographic step with
CC	immobilised enzymes and/or antibodies against inhibitors. Also
CC	included are an isolated nucleic acid molecule encoding an inhibitor
CC	which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
CC	alpha-1-arabino-furanosidase and/or other cellulase, xylan,
CC	arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
CC	nucleic acid, an expression system transformed with the nucleic
CC	acid, a host organism transformed with the nucleic acid, the inhibitory
CC	proteins encoded by the nucleic acids and modulators of the proteins.
CC	A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC	plant or plant materials transformed with the nucleic acid are useful
CC	for the formation of an endoxylanase-inhibitor complex, screening
CC	endoxylanases that are totally, less or not inhibited by the inhibitors,
CC	reducing syruup in refrigerated dough compositions, affecting the
CC	relative affinity and/or relative hydrolysis specificity and/or relative
CC	hydrolysis rate versus water-extractable and/or water-unextractable
CC	arabinoxylans of endoxylanases such as by the formation of an
CC	endoxylanase/inhibitor complex, improving the malting of cereals such as
CC	barley, sorghum and wheat and/or the production of beer, improving the
CC	production and/or quality of baked or extruded cereal products such as
CC	straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC	biscuits, pasta and noodles, animal feed stuff, improving the production
CC	of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
CC	starch separation and production, improving maize processing,
CC	plant disease resistance and nutraceutical and/or pharmaceutical
CC	applications, improving paper and pulp technologies. The present
CC	sequence isa variant of wheat TAxI-I (T. aestivum L endoxylanase

[illegible]

cellulolytic xylanolytic and/or beta-glucanolytic enzymes comprises screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabinofuranosidase and/or other cellulose, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganism, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syringing in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylanases such as by the formation of an endoxylanase/inhibitor complex, improving the malting of cereals such as barley, sorghum and wheat and/or the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten starch separation and production, improving maize processing, plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence is a variant of wheat TAXI-I (T. aestivum L. endoxylanase inhibitor).

Note: Variant amino acids are highlighted in the specification but no wild-type sequence is shown for comparison.

Sequence 381 AA:

Query Match	100.0%	Score 96;	DB 23;	Length 381;
Best Local Similarity	100.0%	Pred. No. 3.6e-07;		
Matches 17; Conservative	100.0%	0; Mismatches 0;	Indels 0;	Gaps 0;
OY	1	KRLGFSRLPHFTGCCGL 17		
Db	365	KRLGFSRLPHFTGCCGL 381		
RESULT 9				
ID	AAU75814			
	AAU75814 standard; Protein; 381 AA.			
XX				
AC	AAU75814;			
XX				
DT	23-APR-2002 (first entry)			
XX				
DE	Wheat L endoxylanase inhibitor, TAXI 1, microheterogenic variant.			
XX				
Wheat; TAXI-1; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;				
KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;				
KM immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;				
KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;				
KM noodle; animal feed; starch separation; maize processing; malting;				
KM plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.				
XX				
OS	Triticum aestivum.			
XX				
Key	Location/Qualifiers			
FH	Misc-difference 95			
FT	/label= Unknown			
FT	Misc-difference 98			
FT	/label= Unknown			
FT	Misc-difference 101			
FT	/label= Unknown			
FT	Misc-difference 110			
FT	/label= Unknown			

FT	Misc-difference	145	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	183	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	232	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	275	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	282	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	333	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference	Unknown	/label= Unknown
PN	WO200198474-A1.		
XX			
PD	27-DEC-2001.		
XX			
PF	21-JUN-2001; 2001WO-BE00106.		
XX			
PR	22-JUN-2000; 2000GB-0015296.		
PR	25-JAN-2001; 2001GB-0002018.		
PR	26-JAN-2001; 2001GB-0002194.		
PR	16-MAR-2001; 2001GB-0006364.		
PR	21-MAY-2001; 2001GB-0012328.		
PA	(LEUV-) LEUVEN RES & DEV.		
XX			
PI	Delcour J, Debysse W, Gebruers K, Goessert H, Fierens K, Robben J;		
PI	Van Campenhout S;		
DR	WPI: 2002-114579/15.		
XX			
PT	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or		
PT	beta-glucanolytic enzymes comprises using endoxylanases during		
PT	screening for inhibition activity or affinity chromatography with		
PT	immobilised enzymes -		
XX			
PS	Claim 127; page 51; 127pp; English.		
XX			
CC	The invention relates to separating and/or isolating inhibitors of		
CC	cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising		
CC	screening the inhibition activity by using two or more enzymes during the		
CC	separation and/or isolation steps that allow to distinguish inhibitors of		
CC	different specificity or by using an affinity chromatographic step with		
CC	immobilised enzymes and/or antibodies against inhibitors. Also		
CC	included are an isolated nucleic acid molecule encoding an inhibitor		
CC	which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylanase,		
CC	alpha-L-arabino-furanosidase and/or other cellulose, xylan,		
CC	arabinofuran or beta-glucan degrading enzymes, a vector comprising the		
CC	nucleic acid, an expression system transformed with the nucleic		
CC	acid, a host organism transformed with the nucleic acid, the inhibitory		
CC	proteins encoded by the nucleic acids and modulators of the proteins.		
CC	A recombinant protein, glycoprotein or polypeptide or microorganisms,		
CC	plant or plant materials transformed with the nucleic acid are useful		
CC	for the formation of an endoxylanase-inhibitor complex, screening		
CC	endoxylanases that are totally, less or not inhibited by the inhibitors,		
CC	reducing synerging in refrigerated dough compositions, affecting the		
CC	relative affinity and/or relative hydrolysis specificity and/or relative		
CC	hydrolysis rate versus water-extractable and/or water-unextractable		
CC	arabinofuran of endoxylanases such as by the formation of an		
CC	endoxylanase/inhibitor complex, improving the malting of cereals such as		
CC	barley, sorghum and wheat and/or the production of beer, improving the		
CC	production and/or quality of baked or extruded cereal products such as		
CC	straight dough, sponge dough, Chorleywood bread, breakfast cereals,		
CC	biscuits, pasta and noodles, animal feed stuff, improving the production		
CC	of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten		
CC	starch separation and production, improving maize processing,		
CC	plant disease resistance and nutraceutical and/or pharmaceutical		
CC	applications, improving paper and pulp technologies. The present		
CC	sequence is a microheterogenic variant of wheat TxAI-I (T. aestivum L		
CC	endoxylanase inhibitor).		
CC	Note: Variant amino acids are highlighted in the specification but		
CC	no wild-type sequence is shown for comparison.		
CC			

```

XX Sequence      381 AA:
SQ
Query Match
Best Local Similarity 100.0%; Score 96; DB 23; Length 381;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRIGFSRLPHFTGCGGL 17
   |||||
Db 365 KRIGFSRLPHFTGCGGL 381

RESULT 10
AAU75829
ID AAU75829 standard; Peptide: 14 AA.
XX
AC AAU75829;
XX
DT 23-APR-2002 (first entry)
XX
DE Wheat xylanase inhibitor TAXI-I PCR product C-terminal sequence.
XX
KW Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
OS Triticum aestivum cultivar Estica.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 11 /label= OTHER
FT /note= "Other- In-frame stop codon"
XX
XX WO200198474-A1.
XX
PD 27-DEC-2001.
XX
PE 21-JUN-2001; 2001WO-BE00106.
XX
PR 22-JUN-2000; 2000GB-0015296.
PR 25-JAN-2001; 2001GB-0002018.
PR 26-JAN-2001; 2001GB-0002194.
PR 16-MAR-2001; 2001GB-0006564.
PR 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES & DEV.
XX
PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
PI Van Campenhout S;
XX
DR WPI; 2002-114579/15.
XX
DR N-PSDB; ABK13691.
XX
PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT beta-glucanolytic enzymes comprises using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
PS Example 11; Fig 21; 127pp; English.
XX
CC The invention relates to separating and/or isolating inhibitors of
CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
CC screening the inhibition activity by using two or more enzymes during the
CC separation and/or isolation steps that allow to distinguish inhibitors of
CC different specificity or by using an affinity chromatographic step with
CC immobilised enzymes and/or antibodies against inhibitors. Also
CC included are an isolated nucleic acid molecule encoding an inhibitor
CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,

```

```

CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
CC nucleic acid, an expression system transformed with the nucleic
CC acid, a host organism transformed with the nucleic acid, the inhibitory
CC proteins encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syripping in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specificity and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC endoxylanase/inhibitor complex, improving the malting of cereals such as
CC barley, sorghum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
CC starch separation and production, improving maize processing,
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence represents the sequence surrounding the C-terminus of wheat
CC TAXI-I (T.aestivum L endoxylanase inhibitor) which was expressed in
CC E. coli.
XX
SQ Sequence      14 AA:
Query Match
Best Local Similarity 63.5%; Score 61; DB 23; Length 14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LPHFTGCGGL 17
   |||||
Db 1 LPHFTGCGGL 10

RESULT 11
AAG24723
ID AAG24723 standard; Protein: 198 AA.
XX
AC AAG24723;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28506.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.

```

PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 16-JUN-1999: 99US-0139454.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144684.
PR 21-JUL-1999: 99US-0144814.

PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145089.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145313.
PR 27-JUL-1999: 99US-0145918.
PR 28-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148665.
PR 16-AUG-1999: 99US-0148684.
PR 17-AUG-1999: 99US-0149368.
PR 18-AUG-1999: 99US-0149175.
PR 20-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154079.
PR 22-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157153.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159294.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.

PR	23-JUL-1999;	99US-0145210;
PR	23-JUL-1999;	99US-0145226
PR	26-JUL-1999;	99US-0145274
PR	27-JUL-1999;	99US-0145913
PR	27-JUL-1999;	99US-0145918
PR	27-JUL-1999;	99US-0145919
PR	28-JUL-1999;	99US-0145951
PR	02-AUG-1999;	99US-0145386
PR	02-AUG-1999;	99US-0146389
PR	03-AUG-1999;	99US-0147004
PR	04-AUG-1999;	99US-0147302
PR	05-AUG-1999;	99US-0147192
PR	05-AUG-1999;	99US-0147260
PR	06-AUG-1999;	99US-0147303
PR	06-AUG-1999;	99US-0147416
PR	09-AUG-1999;	99US-0147493
PR	09-AUG-1999;	99US-0147493
PR	10-AUG-1999;	99US-0148171
PR	11-AUG-1999;	99US-0148319
PR	12-AUG-1999;	99US-0148341
PR	13-AUG-1999;	99US-0148565
PR	13-AUG-1999;	99US-0148684
PR	16-AUG-1999;	99US-0149368
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149426
PR	20-AUG-1999;	99US-0149723
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149929
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150566
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151080
PR	30-AUG-1999;	99US-0151303
PR	31-AUG-1999;	99US-0151438
PR	01-SEP-1999;	99US-0151930
PR	07-SEP-1999;	99US-0152363
PR	10-SEP-1999;	99US-0153700
PR	13-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	16-SEP-1999;	99US-0154039
PR	20-SEP-1999;	99US-0154479
PR	20-SEP-1999;	99US-0155139
PR	23-SEP-1999;	99US-0155486
PR	24-SEP-1999;	99US-0155659
PR	28-SEP-1999;	99US-0155658
PR	28-SEP-1999;	99US-0155648
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	07-OCT-1999;	99US-0157865
PR	08-OCT-1999;	99US-0158209
PR	12-OCT-1999;	99US-0158369
PR	13-OCT-1999;	99US-0159293
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159637
PR	18-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160747
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160776
PR	21-OCT-1999;	99US-0160814
PR	22-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980

PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity	48.4%;	Score 46.5; DB 21; Length 198;
Matches 10; Conservative	43.5%;	Pred. No. 14;
	1;	Mismatches 3; Indels 9; Gaps 1;
OY	2 RLGESRLP-----HFTGCG 15 I: I III II III	
Db	7 RVSFRRLPQSVSLQRHFRLGCG 29	
RESULT 13		
AAG24722	AAG24722 standard; Protein: 252 AA.	
ID	AAG24722	
XX	AAC24722;	
AC	AAC24722;	
XX	AAC24722;	
DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 28505.	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 28505.	
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
PD	06-SEP-2000.	
XX	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
XX	25-FEB-2000; 2000EP-0301439.	
PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 48.4%; Score 46.5; DB 21; Length 252;
Best Local Similarity 43.5%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 3; Indels 9; Caps 1;

QY 2 RLGSRP-----HFGCG 15
I:|||||
61 RVSRRLPGSVSLRRHFLGCG 83

RESULT 14
ID AAG45303 standard; Protein; 252 AA.
AC AAG45303;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56858.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 18-MAY-1999; 990S-0134941.
PR 19-MAY-1999; 990S-0135124.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.

PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 17-JUN-1999; 990S-0139453.
PR 18-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142820.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147792.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

48.4%; Score 46.5; DB 21; Length 252;

Best Local Similarity 43.5%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 3; Indels 9; Gaps 1;
QY 2 RLGRSLP-----HTGCG 15
|:| ||| | |||
Db 61 RVSEFRLPGSVSLQRRHFLGCG 83
RESULT 15
ID AAG24721 standard; Protein; 256 AA.
AC AAG24721;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28504.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 200EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132853.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

```
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145313.
PR 27-JUL-1999; 990S-0145318.
PR 27-JUL-1999; 990S-0145319.
PR 28-JUL-1999; 990S-0145361.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.

PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149176.
PR 20-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 23-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158322.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 48.4%; Score 46.5; DB 21; length 256;
Best Local Similarity 43.5%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 9; Gaps 1;

QY 2 RLGSRP-----HFTGCG 15
DB 65 KVSFRRLPQSVSLQRRHFLGCG 87
```

Tue May 20 18:04:53 2003

us-09-869-155-14.rag

Page 15

Search completed: May 20, 2003, 16:14:38
Job time : 19.4945 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 6.91209 Seconds
(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-14

Perfect score: 96
Sequence: 1 KRLGFSRLPHFTGCGGL 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_73:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46.5	48.4	253	D71446	hypothetical prote
2	45	46.9	290	C71701	branched-chain ami
3	43.5	45.3	438	T52082	alpha galactosyltr
4	43	44.8	206	AG0692	conserved hypothet
5	42	43.8	255	F83058	probable ATP-bind
6	42	43.8	268	T46902	hypothetical prote
7	42	43.8	325	T02117	hypothetical prote
8	42	43.8	472	T48074	hypothetical prote
9	42	43.8	1512	G85090	hypothetical prote
10	42	43.8	1534	S59604	DNA (cytosine-5-)-
11	42	43.8	3097	T28635	glutamate synthase
12	41.5	43.2	376	I49735	hepatocyte nuclear
13	41.5	43.2	581	T05232	hypothetical prote
14	41	42.7	96	T17865	hypothetical prote
15	41	42.7	178	T38624	hypothetical prote
16	41	42.7	211	A82915	ribosomal protein
17	41	42.7	333	S77222	hypothetical prote
18	41	42.7	431	AH0860	probable RNA methyl
19	41	42.7	433	E65060	ycga protein - Esc
20	41	42.7	433	F85929	probable enzyme yg
21	41	42.7	457	T33244	hypothetical prote
22	41	42.7	591	AC3528	extracellular prote
23	41	42.7	729	AC3528	gag polyprotein -
24	41	42.7	758	FOLJSA	probable sugar tra
25	41	42.7	855	JC7731	membrane-bound arg
26	41	42.7	855	JC7775	membrane type-seri
27	41	42.7	958	A70634	probable mmp11 pro
28	41	42.7	1528	D85912	hypothetical prote
29	41	42.7	1569	A65044	hypothetical prote

30	41	42.7	1571	2	C91068	hypothetical prote
31	40.5	42.2	212	2	S73991	ribosomal protein
32	40	41.7	83	2	J50668	myomy protein - ra
33	40	41.7	88	2	A10041	hypothetical prote
34	40	41.7	133	2	G85643	hypothetical prote
35	40	41.7	133	2	E90783	hypothetical prote
36	40	41.7	172	2	E82464	probable acetyltra
37	40	41.7	324	2	JG0163	glucuronyltransfer
38	40	41.7	342	2	B70100	beta-N-acetylhexos
39	40	41.7	382	2	G90751	probable transport
40	40	41.7	382	2	E85615	probable transport
41	40	41.7	382	2	A64829	membrane protein y
42	40	41.7	473	2	S04113	nitrogenase (EC 1.
43	40	41.7	474	2	C35405	nitrogenase (EC 1.
44	40	41.7	528	2	B88989	protein F02C9.2 li
45	40	41.7	564	2	F97668	probable electron

ALIGNMENTS

RESULT 1

D71446
hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C>Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

C:Accession: D71446

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Wellenreger, T.; Pohl, T.M.; Terry, N.; G

avanagh, T.; Hempel, S.; Kotter, P.; Eutlian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen

erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;

C:Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: D71446

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-253 <BEV>

A:Cross-references: GB:T97343; NID:92245073; PID:e327075; PID:92245123

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 48.4%; Score 46.5; DB 2; Length 253;

Best Local Similarity 43.5%; Pred. No. 3.8;

Matches 10; Conservative 1; Mismatches 3; Indels 9; Gaps 1;

QY 2 KRLGFSRLP-----HFTGCG 15

DB 127 RVSFRRLPQSVSLQRRHFLGCG 149

RESULT 2

C71701
branched-chain amino acid aminotransferase (liver) RP428 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: C71701

R:Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9623893

A:Accession: C71701

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-260 <AND>

A:Cross-references: GB:AJ35271; GB:AJ235269; NID:g3868717; PIDN:CAAI4885.1; PID:g386

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: liver; RP428

C:Superfamily: branched-chain-amino-acid aminotransferase

Query Match 46.9%; Score 45; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 7.7;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRIGFSLPHFTGC 14
||| :|||
EB 234 ERLKLAQIEHFTGC 247

RESULT 3

alpha galactosyltransferase (EC 2.4.1.-) [imported] - Trigonella foenum-graecum (fragmen
C:Species: Trigonella foenum-graecum
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52082
R:Stover, C.K.; Pham, X.Q.; Ewin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
Plant, J. 19, 691-697, 1999
A:Title: Molecular characterisation of a membrane-bound galactosyltransferase of plant
A:Reference number: Z25943
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: T52082
A:Molecule type: mRNA
A:Residues: 1-438 <EDW>
A:Cross-references: EMBL:AJ245478; PIDN:CAB52246.1
C:Genetics:
A:Gene: galtran
C:Function:
A:Pathway: galactomannan biosynthesis
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 45.3%; Score 43.5; DB 2; Length 438;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1 KRIGFSLPHFTGC 14
||| :|||
DB 363 KNLGDMKRPRTHTFTGC 379

RESULT 4

AG0692
Conserved hypothetical protein STY1667 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG0692
R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th. T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01912.1; PID:q16502754; GSPDB:GN00176
C:Genetics: STY1667
A:Gene: STY1667
C:Superfamily: hypothetical protein H11687

Query Match 44.8%; Score 43; DB 2; Length 206;
Best Local Similarity 42.9%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

QY 1 KRIGFSLPHFTGC 15
||| :|||
DB 185 KRAGLYAESLPADLPHTLACG 205

RESULT 5

F83058
probable ATP-binding component of ABC transporter PA4706 [imported] - Pseudomonas aer
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83058
R:Stover, C.K.; Pham, X.Q.; Ewin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; PMID:20437337; PMID:10984043
A:Accession: F83058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>
A:Cross-references: GB:AE004884; GB:AE004091; NID:g9950954; PIDN:AAG08092.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4706

Query Match 43.8%; Score 42; DB 2; Length 255;
Best Local Similarity 63.6%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LGFSLPHFTGC 13
||| :|||
DB 96 VGFGRPLPHSSG 106

RESULT 6

T46902
hypothetical protein DKFZp761E10121.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: T46902
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Well, B.; Wiemann, S.
Submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24135
A:Accession: T46902
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-268 <AA>
A:Cross-references: EMBL:AL157417
A:Experimental source: adult amygdala; clone DKFZp761E10121
C:Genetics:
A:Note: DKFZp761E10121.1

Query Match 43.8%; Score 42; DB 2; Length 268;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLGFSLPHFTGC 15
||| :|||
DB 216 KVGFRLPHGWTGCG 229

RESULT 7

T02117
hypothetical protein At2g41020 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T3K9.21
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 16-Feb-2001
C:Accession: T02117; G84836
R:Bounsliev, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, February 1999
A:Description: Arabidopsis thaliana chromosome II BAC T3K9 genomic sequence.
A:Reference number: Z14570
A:Accession: T02117
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <ROU>
A:Cross-references: EMBL:AC004261; NID:g3402695; PID:g3402715
A:Experimental source: cultivar Columbia

A;CROSS-References: GB:NC_001268; NID:g/26/541; P1DN:CAB/8023.1; GSPDB:GN00140

```
C:\Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1995
```

R:Clevidence, D.E.; Overdier, D.G.; Peterson, R.S.; Porcella, A.; Ye, H.; Paulson, K.E.;
Dev. Biol. 166, 195-209, 1994
A>Title: Members of the HNF-3/forkhead family of transcription factors exhibit distinct
A:Reference number: 149735; PMID:95046902; PMID:7958446
A:Accession: 149735
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-376 <RES>
A:Cross-references: GB:L35949; NID:g762833; PIDN:AAA64885.1; PID:g575512
C:Genetics:
A:Gene: HNF-8
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:47-138/Domain: fork head DNA-binding domain homology <FHD>

Query Match 43.2%; Score 41.5; DB 2; Length 376;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 10; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

OY 3 LGFSRLPH---FTGCGGL 17
||| ||| . | |||
Db 160 LGFNHLPTDYGFGCGGL 177

RESULT 13

hypothetical protein F18A5.20 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05232
R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; Maye
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15405
A:Accession: T05232
A:Molecule type: DNA
A:Residues: 1-581 <BEV>
A:Cross-references: EMBL:AL035528
A:Experimental source: cultivar Columbia; BAC clone F18A5
C:Genetics:
A:Map position: 4
A:Introns: 184/1; 568/3
A:Note: F18A5.20

Query Match 43.2%; Score 41.5; DB 2; Length 581;
Best Local Similarity 43.5%; Pred. No. 59;
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

OY 1 KRLGFSRLP-----HFTGCGGL 17
||| ||| | | |||
Db 144 KRLSFRSPYGFKNHYOSCLGL 165

RESULT 14

T17865
hypothetical protein a365L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17865
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T17865
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-96 <GRA>
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96733.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: a365L

Query Match 42.7%; Score 41; DB 2; Length 96;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 LGFSRLPHFTGC 14
||| ||| |||
Db 70 LFFSNLSHTVGC 81

RESULT 15

T38624
hypothetical protein SPAC31G5.07 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T38624
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21722
A:Accession: T38624
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <MCL>
A:Cross-references: EMBL:Z98979; PIDN:CA11691.1; GSPDB:GN00066; SPDB:SPAC31G5.07
A:Experimental source: strain 972h-; cosmid c31G5
C:Genetics:
A:Gene: SPDB:SPAC31G5.07
A:Map position: 1
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC31G5.07

Query Match 42.7%; Score 41; DB 2; Length 178;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 PHFTGCGGL 17
| ||||| ||
Db 14 POFPGCPGL 22

Search completed: May 20, 2003, 16:36:58
Job time : 9.91209 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 4.57692 Seconds

(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-14

Perfect score: 96

Sequence: 1 KRLGSRLLPHTGCCGL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47.5	49.5	873	1 RX_DROME	Q9W2Q1 drosophila
2	45	46.9	290	1 ILVE_RICPR	O05970 rickettsia
3	44	45.8	933	1 PLD2_MOUSE	P97813 mus musculu
4	44	45.8	933	1 PLD2_MOUSE	P70498 rattus norv
5	43	44.8	206	1 RNFG_RAT	O82697 salmone
6	42	43.8	1534	1 DNMI_ARATH	P34861 arabidopsis
7	41.5	43.2	353	1 FXPL_MOUSE	O61080 mus musculu
8	41	42.7	338	1 HIS8_PPRFU	O81225 pyrococcus
9	41	42.7	430	1 RUMA_SALTI	O82446 salmone
10	41	42.7	432	1 RUMA_SALTY	O82446 salmone
11	41	42.7	432	1 RUMA_SALTY	O82446 salmone
12	41	42.7	432	1 RUMA_SALTY	O82446 salmone
13	41	42.7	432	1 RUMA_SALTY	O82446 salmone
14	41	42.7	432	1 RUMA_SALTY	O82446 salmone
15	41	42.7	432	1 RUMA_SALTY	O82446 salmone
16	40.5	42.2	212	1 RLA_MYCPN	P75579 mycoplasma
17	40	41.7	265	1 VNED_AZOPA	O68954 azotobacter
18	40	41.7	279	1 VNED_AZOPA	O68954 azotobacter
19	40	41.7	324	1 B3G2_RAT	O92137 rattus norv
20	40	41.7	382	1 YCAD_ECOLI	P15332 azotobacter
21	40	41.7	472	1 VNED_AZOC	P15332 azotobacter
22	40	41.7	473	1 VNED_AZOC	P15332 azotobacter
23	40	41.7	515	1 Y141_HUMAN	O14154 homo sapien
24	40	41.7	628	1 GIDA_NEIMA	O93441 neisseria m
25	40	41.7	628	1 GIDA_NEIMA	O93441 neisseria m
26	39	40.6	121	1 FLIO_ECOLI	P22886 escherichia
27	39	40.6	156	1 VEG_HPV41	P27355 human papil
28	39	40.6	178	1 TRSF_DROER	O23335 drosophila
29	39	40.6	215	1 RLA_MYCSM	O06114 mycobacteri
30	39	40.6	223	1 RLA_MYCSM	O06114 mycobacteri
31	39	40.6	223	1 RLA_MYCSM	O06114 mycobacteri
32	39	40.6	231	1 RNC_SYNY3	P74368 synchocyst
33	39	40.6	302	1 TFPI_RAT	O02445 rattus norv

34	39	40.6	442	1 VAN1_CANAL	Q00314 candida alb
35	39	40.6	889	1 C122_HUMAN	O9b2g6 homo sapien
36	38.5	40.1	488	1 ARI2_CAERL	O22431 caenorhabdi
37	38	39.6	135	1 RT16_MOUSE	O9cpx7 mus musculu
38	38	39.6	137	1 RT16_MOUSE	O9cpx7 mus musculu
39	38	39.6	137	1 RT16_MOUSE	O9cpx7 mus musculu
40	38	39.6	295	1 MTL5_MOUSE	O9b103 rhizobium l
41	38	39.6	303	1 YE15_YEAST	P40100 saccharomyc
42	38	39.6	304	1 MTH5_HAENI	P45000 haemophilus
43	38	39.6	330	1 MTP2_NEIGO	P08455 neisseria g
44	38	39.6	374	1 MTF7_NEIGO	O96006 neisseria g
45	38	39.6	412	1 PEPT_PASWU	O9cp05 pasteurrella

ALIGNMENTS

RESULT 1	ID	RX_DROME	STANDARD	PRT	873 AA.
AC	Q9W2Q1	O46035			
DT	15-OCT-2001	(Rel. 40, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Retinal homeobox protein Rx (DRX1) (DRX).				
GN	RX OR CG10052.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;				
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Canlon S; TISSUE=Embryo;				
RX	MEDLINE=98151514; PubMed=9482887;				
RA	Eggert T., Hauck B., Hildebrandt N., Gehring W.J., Walldorf U.;				
RT	Isolation of a Drosophila homolog of the vertebrate homeobox gene Rx				
RT	and its possible role in brain and eye development."				
RT	Proc. Natl. Acad. Sci. U.S.A. 95:2343-2348(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Berkeley;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	Georgie R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,				
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriil J.F., Aghayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison T.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merklou G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,				
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,				

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [3]
 RN REVISIONS.
 RC STRAIN-Berkeley:
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celisner S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Krommiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnak F.,
 RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Kungall C.J., Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE-97320497; PubMed-9177348;
 RA Natchers P.H., Grindberg A., Mahon K.A., Jamrich M.;
 RT "The Rx homeobox gene is essential for vertebrate eye development.";
 Nature 387:603-607(1997).
 CC -1- FUNCTION: Appears to function in brain development.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- DEVELOPMENTAL STAGE: Expressed in the procephalic region and in
 CC the clypeolabrum from stage 8 on and later in the brain and the
 CC central nervous system.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY.
 CC "BICOID" SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ223300; CA11241.1; ALT_INIT.
 DR EMBL: AE003452; AA46639.2; ALT_SEQ.
 DR TRANSFAC: T03511; -
 DR FlyBase: FBgn020617; Rx.
 DR InterPro: IPR003654; Homeo_OAR.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 DR PROSITE: PS50803; OAR; 1.
 DR Homeobox: DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 KM
 FT DOMAIN 116 123 OCTAPEPTIDE MOTIF.
 FT DNA_BIND 526 586 HOMEBOX.
 FT DOMAIN 849 862 OAR.
 FT DOMAIN 855 859 HOMEBOX.
 FT DOMAIN 329 333 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 366 369 POLY-PRO.
 FT DOMAIN 437 453 POLY-ASP.
 FT DOMAIN 670 673 POLY-GLN.
 FT DOMAIN 683 691 POLY-PRO.
 FT DOMAIN 753 763 POLY-SER.
 FT DOMAIN 828 834 POLY-PRO.
 FT CONFLICT 4 4 S -> P (IN REF. 1).

FT CONFLICT 136 136 R -> W (IN REF. 1).
 FT CONFLICT 384 384 S -> T (IN REF. 1).
 FT CONFLICT 645 672 PLSLAPGNLTMSLSIAMGHHANCP -> QGARKSDH
 FT CONFLICT 768 768 EOSGGHPPCPQMAA (IN REF. 1).
 FT CONFLICT 873 873 G -> L (IN REF. 1).
 SO SEQUENCE 873 AA: 92897 MW: 87856AD693F6710 CRC64;
 Query Match 49.5%; Score 47.5; DB 1; Length 873;
 Best local Similarity 58.8%; Pred. No. 3.4;
 Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
 QY 2 RLG---FSRLPHFTGCG 15
 DB 593 RLGLHTFOLPHRLGCG 609
 RESULT 2
 ILVE_RICPR
 ID ILVE_RICPR STANDARD; PRT; 290 AA.
 AC 005970;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable branched-chain amino acid aminotransferase (EC 2.6.1.42)
 DE (BCAT).
 GN ILVE OR RP428.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiinae; Rickettsia.
 OX NCBI_TaxId=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Madrid E;
 RX MEDLINE-97419517; PubMed-9274032;
 RA Andersson J.O., Andersson S.G.E.;
 RT "Genomic rearrangements during evolution of the obligate
 RT intracellular parasite *Rickettsia prowazekii* as inferred from an
 RT analysis of 52015 bp nucleotide sequence.";
 RL Microbiology 143:2783-2795(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Madrid E;
 RX MEDLINE-9823893; PubMed-9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sichteritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of *Rickettsia prowazekii* and the origin of
 RT mitochondria.";
 RL Nature 396:133-140(1998).
 CC -1- FUNCTION: Acts on leucine, isoleucine and valine (By similarity).
 CC -1- CATALYTIC ACTIVITY: L-leucine + 2-oxoglutarate = 4-methyl-2-
 CC oxopentanoate + L-glutamate.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: Valine and isoleucine biosynthesis.
 CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y11777; CA172450.1; -
 DR EMBL: AJ235271; CA14885.1; -
 DR HSSP: P00510; IA3G.
 DR InterPro: IPR001544; Aminotran_4.
 DR Pfam: PF01063; aminotran_4; 1.
 DR ProDom: PD001961; Aminotran_4; 1.
 DR PROSITE: PS001770; AA-TRANSFER_CLASS_4; 1.
 DR Transferase: Aminotransferase; Branched-chain amino acid biosynthesis;
 KW

KW Pyridoxal phosphate; Complete proteome.
 FT BINDING 155 155 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 290 AA; 33068 MW; 98374E325350763D CRC64;
 Query Match 46.9%; Score 45; DB 1; Length 290;
 Best Local Similarity 50.0%; Pred. No. 3;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRIGFSLPHFTGCG 14
 DB 234 ERLKAOIEHFTGC 247

RESULT 3
 PLD2_MOUSE STANDARD; PRT; 933 AA.
 AC P97813;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phospholipase D2 (EC 3.1.4.4) (PLD 2) (Choline phosphatase 2)
 GN Phosphatidylcholine-hydrolyzing phospholipase D2 (PLD1C) (mPLD2).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo, and Neonatal brain;
 RA MEDLINE=97199446; PubMed=9395408;
 RA Colley W.C., Sung T.-C., Roll R., Jenco J.M., Hammond S.M.,
 RA Altschuler Y.M., Bar-Sagi D., Morris A.J., Frohman M.A.;
 RT "Phospholipase D2, a distinct phospholipase D isoform with novel
 RT regulatory properties that provokes cytoskeletal reorganization.";
 RT Curr. Biol. 7:191-201(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Redina O.E., Frohman M.A.;
 RT "Organization and alternative splicing of the murine phospholipase D2
 RT gene.";
 RT Biochem. J. 331:845-851(1998).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=97439716; PubMed=9307024;
 RA Colley W.C., Altschuler Y.M., Sue-Ling C.K., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Branch K.D., Tsitka S.E., Bollag R.J.,
 RA Bollag W.B., Frohman M.A.;
 RT "Cloning and expression analysis of murine phospholipase D1.";
 RT Biochem. J. 326:745-753(1997).
 CC -1- FUNCTION: MAY HAVE A ROLE IN SIGNAL-INDUCED CYTOSKELETAL
 CC REGULATION AND/OR ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = choline + a
 CC phosphatide.
 CC -1- ENZYME REGULATION: STIMULATED BY PHOSPHATIDYLINOSITOL 4,5-
 CC BISPHOSPHATE. IS NOT RESPONSIVE TO ADP-RIBOSYLATION FACTOR-1 (ARF-
 CC 1), NOR TO GTP-BINDING PROTEINS. RHO A.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS. ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS IN BRAIN AND LUNG.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN THE HIPPOCAMPUS
 CC AT THE EARLIEST TIME AT WHICH IT IS DEFINED AS A STRUCTURE AND
 CC ALSO IN VENTRICULAR NEURAL CELLS AS WELL AS DIFFERENTIATING
 CC NEURONS OUTSIDE OF THE VENTRICULAR REGION. EXPRESSED DURING
 CC DEVELOPMENT IN LOWER LEVELS IN MESENCHYMAL CELLS DERIVED FROM THE
 CC NEURAL CREST THAT ARE DESTINED TO FORM BONES OF THE SKULL.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE D FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PLD PHOSPHODIESTERASE DOMAINS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U87557; AAC53173.1; -;
 CC DR EMBL: AF052294; AAC24519.1;
 CC DR EMBL: AF052291; AAC24519.1; JOINED.
 CC DR EMBL: AF052293; AAC24519.1; JOINED.
 CC DR EMBL: AF052292; AAC24519.1; JOINED.
 CC MGD: MGI:892877; Pld2.
 CC DR InterPro: IPR001849; PH.
 CC DR InterPro: IPR001736; PLD.
 CC DR InterPro: IPR001683; PX.
 CC DR Pfam: PF00169; PH; 1.
 CC DR Pfam: PF00614; PLDc; 2.
 CC DR Pfam: PF00787; PX; 1.
 CC DR SMART: SM00233; PH; 1.
 CC DR SMART: SM00155; PLDc; 2.
 CC DR SMART: SM00312; PX; 1.
 CC DR PROSITE: PS50003; PH DOMAIN; FALSE_NEG.
 CC DR PROSITE: PS50035; PLD; 2.
 CC DR PROSITE: PS50195; PX; 1.
 CC KW Hydrolyase; Lipid degradation; Membrane; Alternative splicing; Repeat.
 CC FT DOMAIN 65 195
 CC FT DOMAIN 203 311
 CC FT DOMAIN 437 464
 CC FT DOMAIN 751 778
 CC FT DOMAIN 441 788
 CC FT CATALYTIC.
 CC SQ SEQUENCE 933 AA; 106167 MW; BADEIDDFEAC9D CRC64;
 Query Match 45.8%; Score 44; DB 1; Length 933;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KRIGFSLPHFTGCG 15
 DB 211 KRSGHVRVPGFTFCG 225

RESULT 4
 PLD2_RAT STANDARD; PRT; 933 AA.
 ID PLD2_RAT
 AC P70498; 008768;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phospholipase D2 (EC 3.1.4.4) (PLD 2) (Choline phosphatase 2)
 GN Phosphatidylcholine-hydrolyzing phospholipase D2 (PLD1C) (rPLD2).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98194218; PubMed=9533024;
 RA Nakashima S., Matsuda Y., Akao Y., Yoshimura S.-I., Sakai H.,
 RA Hayakawa K., Andoh M., Nozawa Y.;
 RT "Molecular cloning and chromosome mapping of rat phospholipase D
 RT genes, Pld1a, Pld1b and Pld2.";
 RT Cytogenet. Cell Genet. 79:109-113(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=97269050; PubMed=9111050;
 RA Kodaki T., Yamashita S.;
 RT "Cloning, expression, and characterization of a novel phospholipase D
 RT complementary DNA from rat brain.";

```

RL J. Biol. Chem. 272:11408-11413(1997).
RN [3]
RP SEQUENCE OF 445-535 FROM N.A.
RX TISSUE-Glial cell;
RX MEDLINE=96354814; PubMed=8753790;
RA Yoshimura S.-I., Nakashima S., Ohguchi K., Sakai H., Shinoda J.,
RA Sakai N., Nozawa Y.;
RT "Differential mRNA expression of phospholipase D (PLD) isozymes
RT during CAMP-induced differentiation in C6 glioma cells.";
RL Biochem. Biophys. Res. Commun. 225:494-499(1996).
CC -1- FUNCTION: MAY HAVE A ROLE IN SIGNAL-INDUCED CYTOSKELETAL
CC REGULATION AND/OR ENDOCYTOSIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = choline + a
CC phosphatidate.
CC -1- ENZYME REGULATION: STIMULATED BY PHOSPHATIDYLINOSITOL 4,5-
CC BISPHOSPHATE, AND PHOSPHATIDYLETHANOLAMINE. INHIBITED BY
CC PHOSPHATIDYLSELINE AND BY OLEATE. IS NOT RESPONSIVE TO ADP-
CC RIBOSYLATION FACTOR-1 (ARF-1), NOR TO GTP-BINDING PROTEINS: RHO A.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, HEART, KIDNEY,
CC STOMACH, SMALL INTESTINE, COLON, AND TESTIS, AND AT A MUCH LOWER
CC LEVELS IN THYMUS, LIVER AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE D FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PLD PHOSPHODIESTERASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB003172; BAA24078.1; -.
DR EMBL: DB8672; BAA19882.1; -.
DR InterPro: IPR001849; PH.
DR InterPro: IPR001736; PLD.
DR InterPro: IPR001683; PX.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00614; PLDC; 2.
DR Pfam: PF00787; PX; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00155; PLDC; 2.
DR SMART: SM00312; PX; 1.
DR PROSITE: PS50003; PH_DOMAIN; FALSE_NEG.
DR PROSITE: PS50035; PLD; 2.
DR PROSITE: PS50195; PX; 1.
DR Hydrolyase; Lipid degradation; Membrane; Repeat.
KW DOMAIN
FT 65 195
FT DOMAIN
FT 203 311
FT DOMAIN
FT 437 464
FT DOMAIN
FT 751 778
FT DOMAIN
FT 441 788
FT CONFLICT
FT 26 26
FT CONFLICT
FT 125 125
FT CONFLICT
FT 599 599
FT CONFLICT
FT 792 792
FT CONFLICT
FT 817 818
FT CONFLICT
FT 919 924
SQ SEQUENCE 933 AA: 106036 MW: D430843B4D541EEA CRC64;

```

Query Match 45.8%; Score 44; DB 1; Length 933;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KRIGFRLPHFTGCG 15
 DB 211 KRSGGHRVPGFTCGC 225

RESULT 5

```

RNFG_SALTI
ID RNFG_SALTI STANDARD; PRT; 206 AA.
AC 082607;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Electron transport complex protein rnfB.
DE RNFG OR STY1667.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Barthill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Kung'u K.L., Bentley S.D., Holden M.T.G., Sebahina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Inner membrane (Potential).
CC -1- SIMILARITY: BELONGS TO THE RNFG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL627271; CAD01912.1; -.
DR Electron transport; Transmembrane; Inner membrane; Complete proteome.
KW TRANSMEM 10 26
SQ SEQUENCE 206 AA: 22252 MW: D5FBA792801B34A CRC64;

```

Query Match 44.8%; Score 43; DB 1; Length 206;
 Best Local Similarity 42.9%; Pred. No. 4.6;
 Matches 9; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

QY 1 KRIGF-----SRLPHTGCG 15
 DB 185 KRAGLYAESLPADLPHTLTCG 205

```

RESULT 6
ID DNML_ARATH STANDARD; PRT; 1534 AA.
AC P34881;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA (cytosine-5)-methyltransferase Ath1 (EC 2.1.1.37) (DNA
DE methyltransferase Ath1) (DNA Metase Ath1) (M.Ath1).
GN ATH1 OR AT5G49160 OR K2IP3.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```


RC STRAIN=cv. Columbia;
 RX MEDLINE=93281384; PubMed=8389441;
 RA Finnegan E.J., Dennis E.S.;
 RT "Isolation and identification by sequence homology of a putative
 RL cytosine methyltransferase from Arabidopsis thaliana.";
 Nucleic Acids Res. 21:2383-2388(1993).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:379-391(1998).
 CC -1- FUNCTION: Methylates CPG residues.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA = S-adenosyl-L-
 CC homocysteine + DNA containing 5-methylcytosine.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 BAH DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: L10692; AAA32829.1; -
 DR HSSP: O14717; 1G55.
 DR REBASE: 2839; M.Ath1.
 DR InterPro: IPR001025; BAH.
 DR InterPro: IPR001525; C5-DNA_meth.
 DR Pfam: PF00145; DNA_methylase; 1.
 DR Pfam: PF01426; BAH; 2.
 DR PRINTS: PR00105; C5METHTRFRASE.
 DR SMART: SM00439; BAH; 2.
 DR TIGRfam: TIGR00675; dcm; 1.
 DR PROSITE: PS00094; C5_MTASE_1; 1.
 DR PROSITE: PS00095; C5_MTASE_2; 1.
 KW Transferrase; Methyltransferase; DNA-binding; Nuclear protein; Repeat.
 FT DOMAIN 735 869 BAH 1.
 FT DOMAIN 909 1049 BAH 2.
 FT ACT_SITE 1198 1198 BY SIMILARITY.
 FT SEQUENCE 1534 AA; 172430 MW; 23FC944AA7074C5A CRC64;
 SQ
 Query Match 43.8%; Score 42; DB 1; Length 1534;
 Best Local Similarity 47.1%; Pred. No. 50;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=96224034; PubMed=8626802;
 RA Helligvist M., Mahlapuu M., Samuelsson L., Enerbaeck S., Carlsson P.;
 RT "Differential activation of lung-specific genes by two forkhead
 RT proteins, FREAC-1 and FREAC-2.";
 RL J. Biol. Chem. 271:4482-4490(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RX MEDLINE=95046902; PubMed=7958446;
 RA Clevidence D.E., Overdier D.G., Peterson R.S., Porcella A., Ye H.,
 RA Paulson K.E., Costa R.H.;
 RT "Members of the HNF-3/forkhead family of transcription factors
 RT exhibit distinct cellular expression patterns in lung and regulate
 RT the surfactant protein B promoter.";
 RL Dev. Biol. 166:195-209(1994).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LUNG-
 CC SPECIFIC GENES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: LUNG AND PLACENTA.
 CC -1- DOMAIN: ACTIVATION DOMAINS C-TERMINAL OF (AND DISTINCT FROM) THE
 CC FORKHEAD DOMAINS ARE NECESSARY FOR TRANSCRIPTIONAL ACTIVATION (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: U42596; AAC52445.1; -
 DR EMBL: L35949; AAA64885.1; ALT_INIT.
 DR HSSP: O63245; 2HFH.
 DR TRANSFAC: T02461; -
 DR MGD: MGI:1347470; Foxf1a.
 DR InterPro: IPR001766; TF_Fork_head.
 DR Pfam: PF00250; Fork_head; 1.
 DR PRINTS: PR00053; FORKHEAD.
 DR Prodom: PD000425; TF_Fork_head; 1.
 DR SMART: SM00339; FH; 1.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; 1.
 DR PROSITE: PS50039; FORK_HEAD_3; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Activator.
 FT DNA_BIND 22 113 FORK-HEAD.
 FT CONFLICT 7 19 GPTKAKTNAGVR -> PHGOEDQRRRA (IN
 FT REF. 2).
 FT CONFLICT 207 247 AGEYPHDSVPASPLPAGAGGVMEPAHYVSSSAAMP
 FT -> GRGVPAPIROLGARFTAPAPAEWSRTPFPALQOPG
 FT RP (IN REF. 2).
 FT SEQUENCE 353 AA; 37798 MW; 046AED08D1765A69 CRC64;
 SQ
 Query Match 43.2%; Score 41.5; DB 1; Length 353;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 10; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

RESULT 7
 FXFL_MOUSE STANDARD: PRT; 353 AA.
 AC 061080; 061661;
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Forkhead box protein F1 (Forkhead-related protein FKHL5) (Forkhead-
 DE related transcription factor 1) (FREAC-1) (Hepatocyte nuclear factor 3
 DE forkhead homolog 8) (HFH-8).
 GN FOXF1 OR FOXF1A OR FKHL5 OR FREAC1 OR HFH8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

RESULT 8
 HIS8_PYRFU STANDARD: PRT; 338 AA.
 AC 08TH25;
 DT 15-JUN-2002 (Rel. 41; Last sequence update)

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Histidinol-phosphate aminotransferase (EC 2.6.1.9) (Imidazole acetol-
GN HisC OR Pf1665.
OS Pyrococcus furiosus.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-histidinol-phosphate + 2-oxoglutarate -> 3-
CC (Imidazol-4-yl)-2-oxopropyl phosphate + L-glutamate.
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Histidine biosynthesis; seventh step.
CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES. HISTIDINOL-PHOSPHATE AMINOTRANSFERASES
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE010265; AAL81789.1; -
DR PROSITE: PS00599; AA_TRANSF_2; 1.
KW Histidine biosynthesis; Transferase; Aminotransferase;
KW Pyridoxal phosphate; Complete proteome.
FT BINDING 204 204 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 338 AA; 39086 MW; A1781B9B375016C CRC64;

Query Match
Best Local Similarity 42.7%; Score 41; DB 1; Length 338;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRUGSRLPHFT 12
DB 45 KRIEFNRPHT 56

RESULT 9
RUMA_SALTY STANDARD: PRT; 430 AA.
ID RUMA_SALTY Q82446;
AC Q82446;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 23S rRNA (Uracil-5-)-methyltransferase ruma (EC 2.1.1.-) (23S rRNA(M-
GN 5-01939)-methyltransferase).
OS Salmonella typhi.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella

```

```

RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
CC 1939 (M-5-01939) in 23S rRNA (By similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing thymine.
CC -1- SIMILARITY: BELONGS TO THE RNA M50 METHYLTRANSFERASE FAMILY. RUMA
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AL627266; CAD06071.1; -
DR TIGRfams: TIGR00479; Y9CA.1.
DR PROSITE: PS01230; TRMA.1; 1.
DR PROSITE: PS01231; TRMA.2; 1.
KW Transferase; Methyltransferase; RNA processing; Complete proteome.
FT INIT_MER 0 0 BY SIMILARITY.
FT ACT_SITE 387 387 BY SIMILARITY.
SQ SEQUENCE 430 AA; 47745 MW; 535E6AE51BD6376 CRC64;

Query Match
Best Local Similarity 42.7%; Score 41; DB 1; Length 430;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLPHFTGCG 16
DB 79 RCPHFGVCG 88

RESULT 10
RUMA_SALTY STANDARD: PRT; 430 AA.
ID RUMA_SALTY Q82ME1;
AC Q82ME1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 23S rRNA (Uracil-5-)-methyltransferase ruma (EC 2.1.1.-) (23S rRNA(M-
GN 5-01939)-methyltransferase).
OS Salmonella typhimurium.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvany E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
CC 1939 (M-5-01939) in 23S rRNA (By similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing thymine.
CC -1- SIMILARITY: BELONGS TO THE RNA M50 METHYLTRANSFERASE FAMILY. RUMA
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: AE008835; AAL21837.1; -
 DR TIGRFS: TIGR00479; ygcA.1.
 DR PROSITE: PS01230; TRNA_1; 1.
 DR PROSITE: PS01231; TRNA_2; 1.
 DR StyGene: SG77272; ruma.
 KW transferase; Methyltransferase; RNA processing; Complete proteome.
 FT INIT MET 0 BY SIMILARITY.
 FT ACT_SITE 387 387 BY SIMILARITY.
 SQ SEQUENCE 430 AA; 47636 MW; 4C415A92515D28DF CRC64;

Query Match 42.7%; Score 41; DB 1; Length 430;
 Best Local Similarity 70.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLPHFTGCGG 16
 I I I I I
 Db 79 RCPHFGVCGG 88

RESULT 11

RUMA_ECO57 STANDARD; PRT; 432 AA.
 ID RUMA_ECO57
 AC O8XED8;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 23S rRNA (Uracil-5'-methyltransferase ruma (EC 2.1.1.-) (23S rRNA(M-5-01939)-methyltransferase).
 GN RUMA OR Z4100 OR ECS3645.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=83334;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Balthner F.R.;
 RL "Genome sequence of enterohemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
 CC 1939 (M-5-01939) in 23S rRNA (By similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
 CC homocysteine + rRNA containing thymine.
 CC -1- SIMILARITY: BELONGS TO THE RNA M5U METHYLTRANSFERASE FAMILY. RUMA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AE005506; AAG57898.1; -
 DR EMBL: AP002562; BAB37068.1; -
 DR TIGRFS: TIGR00479; ygcA.1.
 DR PROSITE: PS01230; TRNA_1; 1.
 DR PROSITE: PS01231; TRNA_2; 1.
 KW transferase; Methyltransferase; RNA processing; Complete proteome.
 FT INIT MET 0 BY SIMILARITY.
 FT ACT_SITE 388 388 BY SIMILARITY.
 FT CONFLICT 88 88 G->S (IN REF. 2).
 SQ SEQUENCE 432 AA; 48024 MW; 9D7F8234E81F3ADE CRC64;

Query Match 42.7%; Score 41; DB 1; Length 432;
 Best Local Similarity 70.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLPHFTGCGG 16
 I I I I I
 Db 79 RCPHFGVCGG 88

RESULT 12

RUMA_ECOLI STANDARD; PRT; 432 AA.
 ID RUMA_ECOLI
 AC P5135;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 23S rRNA (Uracil-5'-methyltransferase ruma (EC 2.1.1.-) (23S rRNA(M-5-01939)-methyltransferase).
 GN RUMA OR B2785.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Balthner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN (12)
 RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
 RX MEDLINE=21863956; PubMed=11779873;
 RA Agarwalla S., Kealey J.T., Santl D.V., Stroud R.M.;
 RT "Characterization of the 23 S ribosomal RNA m5U1939 methyltransferase
 from Escherichia coli";
 RL J. Biol. Chem. 277:8835-8840(2002).
 CC -1- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
 CC 1939 (M-5-01939) in 23S rRNA.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
 CC homocysteine + rRNA containing thymine.
 CC -1- MISCELLANEOUS: Two bound iron atoms are present in each monomer of
 CC ruma. However, the proposed mechanism does not involve metal ions
 CC in the reaction, so it seems unlikely that iron is required for
 CC the methyltransferase activity.
 CC -1- SIMILARITY: BELONGS TO THE RNA M5U METHYLTRANSFERASE FAMILY. RUMA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U29580; AAA69295.1; -
 DR EMBL: AE000362; AAC75827.1; -

DR Ecogène; EG11247; ruma.
 DR InterPro: IPR002792; DOF90.
 DR InterPro: IPR000051; SAM_bind.
 DR InterPro: IPR001566; TRMA_1.
 DR Pfam: PF01938; TRAM_1.
 DR TIGRFAMs: TIGR00479; y9ca.1.
 DR PROSITE: PS01230; TRMA_1; 1.
 DR PROSITE: PS01231; TRMA_2; 1.
 KM Transferase; Methyltransferase; rRNA processing; Complete proteome.
 FT INT_MET 0
 FT ACT_SITE 388 388 BY SIMILARITY
 SO SEQUENCE 432 AA; 47921 MW; DB8CB9F9E5FAB3 CRC64;
 Query Match 42.7%; Score 41; DB 1; Length 432;
 Best Local Similarity 70.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 RLPHFTGCGG 16
 Db 79 RCPHFGVCGG 88
 RESULT 13
 GAG_Srv1 STANDARD; PRT; 658 AA.
 ID GAG_Srv1
 AC P04022;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GAG polyprotein (Core polyprotein) [Contains: Core protein P10; Core phosphoprotein P18; Core protein P12; Core protein P27; Core protein P14; Core protein P4].
 GN GAG.
 OS Simian retrovirus SRV-1.
 OC Viruses; Retroid viruses; Retroviridae.
 OC NCBI_Taxid=11942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86151668; PubMed=1006247;
 RA Power M.D., Marx P.A., Bryant M.L., Gardner M.B., Barr P.J.,
 RA Luciw P.A.;
 RT "Nucleotide sequence of SRV-1, a type D simian acquired immune deficiency syndrome retrovirus.";
 RL Science 231:1567-1572(1986).
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M1841; AAA47730.1; -
 DR PIR: A03951; FOLISA.
 DR InterPro: IPR003322; Gag_p10.
 DR InterPro: IPR000721; Gag_p24.
 DR InterPro: IPR001878; ZnF_CCHC.
 DR Pfam: PF00098; zf-CCHC; 2.
 DR Pfam: PF00607; Gag_p24; 1.
 DR Pfam: PF02337; Gag_p10; 1.
 DR PRINTS: PR00939; C2HCZNFINGER.
 DR SMART: SM00343; ZnF_C2HC; 2.
 DR PROSITE: PS50158; ZF_CCHC; 1.
 KM Core protein: Polyprotein; Phosphorylation; Zinc-finger; Repeat.
 FT CHAIN 1 162
 FT CHAIN 163 217 CORE PROTEIN P10.
 FT CHAIN 218 300 CORE PHOSPHOPROTEIN P18.
 FT CHAIN 301 526 CORE PROTEIN P12.
 FT CHAIN 527 622 CORE PROTEIN P27.
 FT CHAIN 623 658 CORE PROTEIN P14.
 FT CHAIN 658 658 CORE PROTEIN P4.
 FT ZN_FING 548 565 CCHC-TYPE 1.

FT ZN_FING 577 594 CCHC-TYPE 2.
 SO SEQUENCE 658 AA; 73195 MW; 60929C787AF6923A CRC64;
 Query Match 42.7%; Score 41; DB 1; Length 658;
 Best Local Similarity 70.0%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 GFSRLPHFTG 13
 Db 268 GFSRLPHWPG 277
 RESULT 14
 MML1_MYCTU STANDARD; PRT; 958 AA.
 ID MML1_MYCTU
 AC P95211;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative membrane protein mmpL1.
 GN MmpL1 OR RV0402C OR MT0412 OR MTCY04D9.15C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv.
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE MmpL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z84725; CAB06593.1; -
 DR EMBL: AE006945; AAK44636.1; -
 DR TIGR: MT0412; -
 DR TubercuList; RV0402c; -
 DR InterPro: IPR004707; ActII.
 DR InterPro: IPR004869; MmpL.
 DR Pfam: PF03176; MmpL; 2.
 DR TIGRFAMs: TIGR00833; actII; 1.
 KM Hypothetical protein: Transmembrane; Complete proteome.
 FT TRANSMEM 19 39
 FT TRANSMEM 192 212 POTENTIAL.

```

FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 252 272 POTENTIAL.
FT TRANSMEM 295 315 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 377 397 POTENTIAL.
FT TRANSMEM 762 782 POTENTIAL.
FT TRANSMEM 791 811 POTENTIAL.
FT TRANSMEM 814 834 POTENTIAL.
FT TRANSMEM 868 888 POTENTIAL.
FT TRANSMEM 890 910 POTENTIAL.
SQ SEQUENCE 958 AA; 104838 MW; 74FC08CF54C3A14 CRC64;

Query Match
Best Local Similarity 42.7%; Score 41; DB 1; Length 958;
Pred. No. 46;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGFSRLPHF 11
111:111:1
DB 313 LGFSRLPHY 321

```

RESULT 15

```

YPIA_ECOLI
ID YPIA_ECOLI STANDARD; PRT; 1569 AA.
AC P52143; P76610; P77017; P77019;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical outer membrane protein ypiA.
GN YPIA OR B2647.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satch Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50,0-68.8 min on the linkage map and
RT analysis of its sequence features."
RL DNA Res. 4:91-113(1997).
CC -1- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC -1- SIMILARITY: STRONG, TO BORDETELLA PERTACTIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U36840; AAA79815.1; ALT_SEQ.
DR EMBL: AE000350; AAC75695.1; -.
DR EMBL: D90889; BAA16514.1; ALT_INIT.
DR EMBL: D90890; BAA16518.1; ALT_INIT.

```

```

DR EcoGene; EG13213; ypiA.
DR InterPro; IPR004899; Pertactin_sup.
DR Pfam; PF03212; Pertactin; 1.
KW Hypothetical protein; Outer membrane; Complete proteome.
SQ SEQUENCE 1569 AA; 162774 MW; 0D006A9A32154596 CRC64;

Query Match
Best Local Similarity 42.7%; Score 41; DB 1; Length 1569;
Pred. No. 76;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 GFSRLPHFTGCG 16
111:111111
DB 508 GYNKLSHFTTGG 520

```

Search completed: May 20, 2003, 16:35:37
 Job time : 6.57692 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time: 21.0165 seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-14
Perfect score: 96
Sequence: 1 KRIGFSRLPFTGCGSL 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_protist:*
11: sp_virus:*
12: sp_virus:*
13: sp_virus:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	61.5	419	10	08S1V0
2	48	50.0	187	2	08RKC9
3	46.5	48.4	198	10	094K29
4	46.5	48.4	253	10	023605
5	46	47.9	505	3	08X0V8
6	46	47.9	602	13	0902B0
7	45	46.9	630	4	096AD9
8	44	45.8	361	11	09CUE9
9	44	45.8	391	10	09FW53
10	44	45.8	719	5	09N8Z8
11	43.5	45.3	438	10	09S756
12	43	44.8	206	16	08ZPM4
13	43	44.8	378	10	0943A3
14	43	44.8	410	6	097645
15	43	44.8	484	2	093548
16	42	43.8	203	10	09C7D9

17	42	43.8	255	16	068877	068877 pseudomonas
18	42	43.8	268	4	09NS99	09NS99 homo sapien
19	42	43.8	324	10	09M7E7	09M7E7 arabidopsis
20	42	43.8	325	10	080682	080682 arabidopsis
21	42	43.8	325	10	09LH52	09LH52 arabidopsis
22	42	43.8	325	10	093X21	093X21 arabidopsis
23	42	43.8	371	16	08Y1X4	08Y1X4 ralsstoma s
24	42	43.8	440	16	092L2C	092L2C rhizobium m
25	42	43.8	472	10	09LZ11	09LZ11 arabidopsis
26	42	43.8	533	16	09KX21	09KX21 streptomyc
27	42	43.8	1444	5	093XW0	093XW0 caenorhabdi
28	42	43.8	1512	10	09M0S8	09M0S8 arabidopsis
29	42	43.8	3097	5	061143	061143 plasmodium
30	41.5	43.2	581	10	09SVP9	09SVP9 arabidopsis
31	41	42.7	96	12	084679	084679 parametium
32	41	42.7	178	3	014107	014107 schizosacch
33	41	42.7	211	16	09P0Q9	09P0Q9 ureaplasma
34	41	42.7	264	5	090527	090527 anopheles s
35	41	42.7	309	5	095XV7	095XV7 caenorhabdi
36	41	42.7	333	16	P73516	P73516 synchocyst
37	41	42.7	338	17	08TH25	08TH25 pyrococcus
38	41	42.7	374	10	0942D5	0942D5 oryza sativ
39	41	42.7	436	5	09N9R6	09N9R6 leishmania
40	41	42.7	457	5	017032	017032 caenorhabdi
41	41	42.7	465	5	09Y029	09Y029 drosophila
42	41	42.7	534	5	P92031	P92031 drosophila
43	41	42.7	534	5	09VJ37	09VJ37 drosophila
44	41	42.7	591	16	08YDM6	08YDM6 brucella me
45	41	42.7	595	10	094AM9	094AM9 arabidopsis

ALIGNMENTS

RESULT 1
ID 08S1V0 PRELIMINARY; PRT: 419 AA.
AC 08S1V0:
DT 01-JUN-2002 (TREMREL_21, Created)
DT 01-JUN-2002 (TREMREL_21, Last sequence update)
DE 01-JUN-2002 (TREMREL_21, Last annotation update)
DT Putative dermal glycoprotein.
GN P0504E02.9.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriophytidae; Oryzoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0504E02.9";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89708.1; -
SQ SEQUENCE 419 AA; 43786 MW; 792B5F5A4F3F3A8 CRC64;
Query Match 61.5%; Score 59; DB 10; Length 419;
Best Local Similarity 71.4%; Pred. No. 0.13;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 2
ID 08RKC9 PRELIMINARY; PRT: 187 AA.
AC 08RKC9:
DT 01-JUN-2002 (TREMREL_21, Created)
DT 01-JUN-2002 (TREMREL_21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE VME protein.
GN VME.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCB1_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3937;
RA Reverchon S.;
RT "vfm genes of Erwinia chrysanthemi modulate the synthesis of multiple
virulence factors."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ438188; CAD27336.1; -
SO SEQUENCE 187 AA; 21903 MW; 9F85CF180AC8EA06 CRC64;

Query Match 50.0%; Score 48; DB 2; Length 187;
Best Local Similarity 72.7%; Pred. No. 4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GFSRLPHTGC 14
DB 119 GFSRLPHTGC 129

RESULT 3
ID Q94K29 PRELIMINARY; PRT; 198 AA.
AC Q94K29;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Putative sperm protein homolog.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Ito S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
GA Goldsmith A.D., Lee J.M., Quach H.L., Tang C.C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shim P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene d14860W."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinozaki K., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RT "Full length cDNA of gene d14860W."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF370345; AAK44160.1; -
DR EMBL: AY062975; AAL34149.1; -
DR InterPro: IPR005340; UPF0083.1.
DR Pfam: PF03654; UPF0083.1.
SO SEQUENCE 198 AA; 22694 MW; 763B9B8721884E40 CRC64;

Query Match 48.4%; Score 46.5; DB 10; Length 198;
Best Local Similarity 43.5%; Pred. No. 7.4;
Matches 10; Conservative 1; Mismatches 3; Indels 9; Gaps 1;

QY 2 RLGSRLP-----HFTGCG 15
DB 127 RVSEFRLPGSVSLQRRHFLCG 149

RESULT 4
ID Q23605 PRELIMINARY; PRT; 253 AA.
AC Q23605;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE DNA Chromosome 4, ESSA I CONTFG fragment NO. 8 (Sperm protein
homolog).
GN ATG47650.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Slekema W., Drost L.,
RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
RA Medler E., Wambutt R., Wetzenecker T., Pohl T.M., Terry N.,
RA Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
RA Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,
RA Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,
RA Vouklatou E., Milioni D., Hatzopoulos P., Pitaravadi E., Obermaier B.,
RA Halbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
RA Palmer K., Benes V., Reichman S., Ansoorge W., Cooke R., Berger C.,
RA Delsen M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,
RA Schueller C., Chalwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
Arabidopsis thaliana."
RL Nature 391:485-488(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z97343; CAB10545.1; -
DR EMBL: AL161546; CAB8768.1; -
SO SEQUENCE 253 AA; 28598 MW; D76F06B21657B9C CRC64;

Query Match 48.4%; Score 46.5; DB 10; Length 253;
Best Local Similarity 43.5%; Pred. No. 9.5;
Matches 10; Conservative 1; Mismatches 3; Indels 9; Gaps 1;

QY 2 RLGSRLP-----HFTGCG 15
DB 127 RVSEFRLPGSVSLQRRHFLCG 149

RESULT 5
ID Q8X0V8 PRELIMINARY; PRT; 505 AA.
AC Q8X0V8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Hypothetical 55.6 kDa protein.
GN 123A4.340.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCB1_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL670009; CAD21385.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 505 AA; 55613 MW; 4E88DDBA4238E7 CRC64;

Query Match 47.9%; Score 46; DB 3; Length 505;
 Best Local Similarity 62.5%; Pred. No. 23;
 Matches 10; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

OY 3 LGFSRLPHFTGC--CG 16
 DB 229 LGFSRLPHFTGCAPDGC 244

RESULT 6

O90ZB0 PRELIMINARY; PRT; 602 AA.
 AC O90ZB0:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Growth hormone receptor.
 GN GHR.
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21297186; PubMed-11404009;
 RA Lee L.T.O., Nong G., Chan Y.H., Tse D.L.Y., Cheng C.H.K.;
 RT "Molecular cloning of a teleost growth hormone receptor and its
 RT functional interaction with human growth hormone.";
 RL Gene 270:121-129(2001).
 DR EMBL; AF293417; AAK60495.1; -
 DR InterPro: IPR002996; CRIA.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003528; Hemtopoptn_L_F1.
 DR SMART; SMO0060; FN3; 1.
 DR PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
 DR Receptor.
 KW SEQUENCE 602 AA; 67262 MW; 4D9F66821C1A029E CRC64;

Query Match 47.9%; Score 46; DB 13; Length 602;
 Best Local Similarity 88.9%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 SRPHFTGC 14
 DB 33 SRPHFTGC 41

RESULT 7

O96AD9 PRELIMINARY; PRT; 630 AA.
 AC O96AD9:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 66.8 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE;
 RA Strausberg R.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC017261; AAH17261.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 630 AA; 66805 MW; FFDD6F5F020FBC80 CRC64;

Query Match 46.9%; Score 45; DB 4; Length 630;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 8 LPHFTGCG 16
 DB 135 LPHFTGCG 143

RESULT 8

O9CUE9 PRELIMINARY; PRT; 361 AA.
 AC O9CUE9:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 4930597B14R1k protein (Fragment).
 GN 4930597B14R1k.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RA MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kono S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasikawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuenli P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Mordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK016403; BAB30217.1; -
 DR HSSP; P11940; ICVU.
 DR MGD; MGI:1922638; 4930597B14R1k.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR SMART; SMO0360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 FT NON_TER 361 361
 FT 1 1
 SQ SEQUENCE 361 AA; 39350 MW; DE4625D464C6FE30 CRC64;

Query Match 45.8%; Score 44; DB 11; Length 361;
 Best Local Similarity 64.3%; Pred. No. 35;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 LGFSRLPHFTGCG 16
 DB 18 LGFSRLPHFTGCG 31

RESULT 9

Q9FW53 PRELIMINARY; PRT; 391 AA.
 ID Q9FW53
 AC Q9FW53
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Putative protoporphyryinogen oxidase.
 GN OSJNB0094K03.11.
 OS Oryza sativa (Rice).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae;
 OC Eurythidaceae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNB0094K03 genomic sequence."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC069145; AAC16857.1;
 SO SEQUENCE 391 AA; 41019 MW; 7B7FA25F21931B39 CRC64;

Query Match 45.8%; Score 44; DB 10; Length 391;
 Best Local Similarity 56.2%; Pred. No. 38;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 LRGFSRLPHTGCGGL 17
 DB 358 RRGRRPRRPGCGGGL 373

RESULT 10
 Q9NR28 PRELIMINARY; PRT; 719 AA.
 ID Q9NR28
 AC Q9NR28;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 79.9 kDa protein.
 GN CHRI.79.
 OS Trypanosoma brucei.
 OC Eukaryota: Euglenozoa: Kinetoplastida: Trypanosomatidae: Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TRE0927;
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
 RA Leonard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
 RA Gerard C., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL355782; CAB95385.1;
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000433; ZnF_ZZ.
 DR SMART: SM00291; ZnF_ZZ; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS01357; ZF_ZZ_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 719 AA; 79901 MW; DIEDD08BD2587AED CRC64;

Query Match 45.8%; Score 44; DB 5; Length 719;
 Best Local Similarity 88.9%; Pred. No. 69;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRGFSRLPHT 11
 DB 352 LRGFSRLPYF 360

RESULT 11
 Q9ST55

ID Q9ST56 PRELIMINARY; PRT; 438 AA.
 AC Q9ST56;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Alpha galactosyltransferase (Fragment).
 GN GALTRAN.
 OS Trigonella foenum-graecum.
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC eurosids I: Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Trigonella.
 OX NCBI_TaxID=78534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20040039; PubMed=10571854;
 RA Edwards M.E., Dickson C.A., Chengappa S., Sidebottom C.M.,
 RA Gidley M.J., Reid G.;
 RT "Molecular Characterisation of a membrane-bound galactosyltransferase
 of plant cell wall matrix biosynthesis."
 RL Plant J. 19:691-697(1999).
 DR EMBL: AJ245478; CAB52246.1;
 KW Glycosyltransferase; Transferase.
 FT NON_TER
 SO SEQUENCE 438 AA; 51282 MW; E6E35B71AD21848A CRC64;

Query Match 45.3%; Score 43; DB 10; Length 438;
 Best Local Similarity 58.8%; Pred. No. 51;
 Matches 10; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

OY 1 KRLGFSRLP--HPTGC 14
 DB 363 KNLGDMRRPFTHTGCG 379

RESULT 12
 Q8ZPM4 PRELIMINARY; PRT; 206 AA.
 ID Q8ZPM4
 AC Q8ZPM4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative Na+-transporting NADH:ubiquinone oxidoreductase gamma
 DE subunit.
 GN YDGP OR STM1455.
 OS Salmonella typhimurium.
 OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SCG31412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanders K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portvik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grevell N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2."
 RL Nature 413:852-856(2001).
 DR EMBL: AE008763; AAL20377.1;
 KW Ubiquinone; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 206 AA; 22238 MW; 5E6BB49636F05C54 CRC64;

Query Match 44.8%; Score 43; DB 16; Length 206;
 Best Local Similarity 42.9%; Pred. No. 29;
 Matches 9; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

OY 1 KRLGFSRLPHTGCG 15
 DB 185 KRLGFSRLPHTGCG 205

RESULT 15		
Q93548		
ID Q93548	PRELIMINARY;	PRT; 484 AA.
AC Q93548;		
DT 01-DEC-2001	(Tremblurel, 19, Created)	
DT 01-DEC-2001	(Tremblurel, 19, Last sequence update)	

```
Search completed: May 20, 2003, 16:43:10
Job time : 24.0165 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 7.84615 seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-14

Perfect score: 96

Sequence: 1 KRIGFSRLPHFTGCGCL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2-6/ptodata/1/1aa/5A_COMB.pep:*

2: /cgn2-6/ptodata/1/1aa/5B_COMB.pep:*

3: /cgn2-6/ptodata/1/1aa/6A_COMB.pep:*

4: /cgn2-6/ptodata/1/1aa/6B_COMB.pep:*

5: /cgn2-6/ptodata/1/1aa/PCtUS_COMB.pep:*

6: /cgn2-6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
perfect greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	45.8	932	3	US-08-968-752B-6
2	44	45.8	932	4	US-09-536-224-6
3	44	44.8	163	4	US-09-562-737-76
4	42	43.8	163	4	US-09-562-737-71
5	41	42.7	163	4	US-09-562-737-72
6	41	42.7	163	4	US-09-562-737-75
7	41	42.7	729	4	US-09-291-922-28
8	40	41.7	163	4	US-09-562-737-79
9	40	41.7	902	4	US-09-644-600-10
10	39	40.6	64	4	US-09-288-143-155
11	38	39.6	40	4	US-09-046-894-39
12	38	39.6	76	1	US-08-702-344-6
13	38	39.6	335	4	US-09-797-906-2
14	38	39.6	336	4	US-09-141-206-7
15	38	39.6	499	2	US-08-820-170A-40
16	38	39.6	499	3	US-09-055-699-40
17	38	39.6	499	4	US-09-273-565-40
18	38	39.6	499	4	US-09-565-538-40
19	38	39.6	499	4	US-09-661-468-40
20	38	39.6	557	4	US-09-138-277C-3
21	38	39.6	698	4	US-09-138-277C-1
22	38	39.6	699	4	US-08-274-121B-6
23	38	39.6	737	4	US-09-291-922-8
24	38	39.6	855	2	US-09-027-337-2
25	38	39.6	855	4	US-09-644-600-2
26	38	39.6	922	4	US-09-141-206-6
27	38	39.6	933	4	US-09-141-206-6

28	38	39.6	933	4	US-09-107-149-17	Sequence 17, Appl
29	38	39.6	1841	2	US-08-804-227C-6	Sequence 6, Appl
30	38	39.6	5215	4	US-09-105-537-2	Sequence 2, Appl
31	37.5	39.1	414	4	US-09-134-001C-5528	Sequence 5528, Ap
32	37	38.5	2071	4	US-09-415-522-6	Sequence 6, Appl
33	36.5	38.0	1151	4	US-09-177-165A-31	Sequence 31, Appl
34	36	37.5	31	3	US-09-100-414B-89	Sequence 89, Appl
35	36	37.5	31	4	US-09-303-323-89	Sequence 89, Appl
36	36	37.5	163	4	US-09-562-737-78	Sequence 78, Appl
37	36	37.5	163	4	US-09-562-737-80	Sequence 80, Appl
38	36	37.5	237	1	US-08-750-532-18	Sequence 18, Appl
39	36	37.5	276	4	US-09-134-001C-3456	Sequence 3456, Ap
40	36	37.5	392	2	US-08-886-152-1	Sequence 1, Appl
41	36	37.5	392	2	US-08-886-152-3	Sequence 3, Appl
42	36	37.5	392	4	US-09-196-232-1	Sequence 1, Appl
43	36	37.5	392	4	US-09-196-232-3	Sequence 3, Appl
44	36	37.5	579	1	US-08-126-564A-31	Sequence 31, Appl
45	36	37.5	579	5	PCT-US94-09143-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-08-968-752B-6

Sequence 6, Application US/08968752B

Patent No. 6043073

GENERAL INFORMATION:

APPLICANT: Frohman, Michael A.

APPLICANT: Morris, Andrew

TITLE OF INVENTION: No. 6043073el Phospholipase D Polypeptide and

TITLE OF INVENTION: DNA Sequences

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: ONYX Pharmaceuticals, Inc.

STREET: 3031 Research Drive

CITY: Richmond

STATE: California

COUNTRY: USA

ZIP: 94806

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/968,752B

FILING DATE: 13-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/025,469

FILING DATE: 05-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Giotta, Gregory J.

REGISTRATION NUMBER: 32,028

REFERENCE/DOCKET NUMBER: ONYX2004

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-222-9700

TELEFAX: 510-222-9758

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 932 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-968-752B-6

Query Match

Best Local Similarity 45.8%; Score 44; DB 3; Length 932;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

1 KRIGFSRLPHFTGCG 15

|| | | | | | | |

DE 211 KRSGGHRVPGFTFCG 225

RESULT 2

US-09-536-224-6
Sequence 6, Application US/09536224
Patent No. 6379665
GENERAL INFORMATION:
APPLICANT: Frohman, Michael A.
TITLE OF INVENTION: No. 6379665el Phospholipase D Polypeptide and
DNA Sequences
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: ONYX Pharmaceuticals, Inc.
STREET: 3031 Research Drive
CITY: Richmond
STATE: California
COUNTRY: USA
ZIP: 94806
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,224
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/968,752
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giotta, Gregory J.
REGISTRATION NUMBER: 32,028
REFERENCE/DOCKET NUMBER: ONYX2004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-222-9700
TELEFAX: 510-222-9758
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-536-224-6

Query Match 45.8%; Score 44; DB 4; Length 932;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KRGSRLPHFTCG 15
DB 211 KRSGGHRVPGFTFCG 225

RESULT 3
US-09-562-737-76
Sequence 76, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 76
LENGTH: 163
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Sequence
US-09-562-737-76

Query Match 44.8%; Score 43; DB 4; Length 163;
Best Local Similarity 66.7%; Pred. No. 8.5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRGSRLPHFT 12
DB 84 KMGFERLEHFS 95

RESULT 4

US-09-562-737-71
Sequence 71, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 71
LENGTH: 163
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-71

Query Match 43.8%; Score 42; DB 4; Length 163;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KRGSRLPHFTG 13
DB 84 KMGFERLEHFSG 96

RESULT 5

US-09-562-737-72
Sequence 72, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 72
LENGTH: 163
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-72

Query Match 42.7%; Score 41; DB 4; Length 163;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KRGSRLPHFTG 13
DB 84 KMGFERLEHFSG 96

Db 84 KMGFGRLNHFSG 96

RESULT 6

US-09-562-737-75
; Sequence 75, Application US/09562737
; Patent No. 6428967

GENERAL INFORMATION:

APPLICANT: Herz, Joachim
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 75
LENGTH: 163

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-562-737-75

Query Match Best Local Similarity 42.7%; Score 41; DB 4; Length 163;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KRLGFSRLPHFTG 13

Db 84 KMGFGRLNHFSG 96

RESULT 7

US-09-291-922-29

; Sequence 29, Application US/09291922
; Patent No. 6383776

GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 29
LENGTH: 729
TYPE: PRT
ORGANISM: Arabidopsis thaliana

US-09-291-922-29

Query Match Best Local Similarity 42.7%; Score 41; DB 4; Length 729;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 LPHFTGCGGL 17

Db 136 LPHFTGCGGL 145

RESULT 8
US-09-562-737-79

; Sequence 79, Application US/09562737
; Patent No. 6428967

GENERAL INFORMATION:

APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael

TITLE OF INVENTION: LDL Receptor Signaling Pathways

FILE REFERENCE: UTSW0708

CURRENT APPLICATION NUMBER: US/09/562,737

CURRENT FILING DATE: 2000-05-01

NUMBER OF SEQ ID NOS: 132

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 79

LENGTH: 163

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-562-737-79

Query Match Best Local Similarity 41.7%; Score 40; DB 4; Length 163;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KRLGFSRLPHFTG 13

Db 84 KMGFGRLNHFSG 96

RESULT 9

US-09-644-600-10

; Sequence 10, Application US/09644600
; Patent No. 6451500

GENERAL INFORMATION:

APPLICANT: O'Brien, Timothy J.
APPLICANT: Tanimoto, Hirotooshi
TITLE OF INVENTION: Overexpressed in Carcinomas
FILE REFERENCE: D6064CIP/D
CURRENT APPLICATION NUMBER: US/09/644,600
CURRENT FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/421,213
PRIOR FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: 09/027,337
PRIOR FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 10
LENGTH: 902
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: Epithin

US-09-644-600-10

Query Match Best Local Similarity 41.7%; Score 40; DB 4; Length 902;
Matches 8; Conservative 4; Mismatches 5; Indels 4; Gaps 1;

OY 1 KRLGFSRLPHFTG 17

Db 323 RHLGFEATFQPLPKMSSCGV 343

RESULT 10

US-09-288-143-155

; Sequence 155, Application US/09288143
; Patent No. 6433139

GENERAL INFORMATION:

APPLICANT: Brewer et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: P2018P1
CURRENT APPLICATION NUMBER: US/09/288,143
CURRENT FILING DATE: 1999-04-08
EARLIER APPLICATION NUMBER: PCT/US98/21142
EARLIER FILING DATE: 1998-10-08
EARLIER APPLICATION NUMBER: 60/061,463
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,529

EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/071,498
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,527
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,536
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,532
NUMBER OF SEQ ID NOS: 219
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 155
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
US-09-288-143-155

Query Match 40.6%; Score 39; DB 4; Length 64;
Best Local Similarity 47.6%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 KRGSRLPHFT---GCCGL 17
1 :||| ||:| |||
16 KMGSIFFHWIWLIEGL 36

De 16 KMGSIFFHWIWLIEGL 36

RESULT 11
US-09-046-894-39
Sequence 39, Application US/09046894
Patent No. 6190857
GENERAL INFORMATION:
APPLICANT: Ralph, David
APPLICANT: An, Gang
APPLICANT: O'Hara, Mark S.
APPLICANT: Velti, Robert
TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
TITLE OF INVENTION: PROFILES IN PERIPHERAL LEUCOCYTES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,894
FILING DATE: Concurrently Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,576
FILING DATE: 24-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Nakashima, Richard A.
REGISTRATION NUMBER: P-42,023
REFERENCE/DOCKET NUMBER: UROC:014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear
US-09-046-894-39

Query Match 39.6%; Score 38; DB 4; Length 40;

Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 PHFTGGC 15
11 |||||
11 PHFTGGC 17

Db 11 PHFTGGC 17

RESULT 12
US-08-702-344-6
Sequence 6, Application US/08702344
Patent No. 5723315
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavallee, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,344
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-702-344-6

Query Match 39.6%; Score 38; DB 1; Length 76;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GFSRLPHFTG 13
11111111
10 GFSYLNHFSG 19

Db 10 GFSYLNHFSG 19

RESULT 13
US-09-797-906-2
Sequence 2, Application US/09797906
Patent No. 6329188
GENERAL INFORMATION:
APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIPRANCESCO, Ellen M. BEASLEY
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: C0001151CIP

;; CURRENT APPLICATION NUMBER: US/09/797,906
;; CURRENT FILING DATE: 2001-03-05
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 2
;; LENGTH: 335
;; TYPE: PRT
;; ORGANISM: Human
US-09-797-906-2

Query Match 39.6%; Score 38; DB 4; Length 335;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 GFGRLPHTGCG 15
DB 242 GFVCPHFVGHG 253

RESULT 14
US-09-141-206-7
; Sequence 7, Application US/09141206
; Patent No. 6187559
; GENERAL INFORMATION:
; APPLICANT: Steed, Paul M.
; APPLICANT: Lasala, Daniel J.
; TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
; Patent No. 6187559
; FILE REFERENCE: 4-30148/P1/CGC1954/R
; CURRENT APPLICATION NUMBER: US/09/141,206
; CURRENT FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/057,802
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Human
US-09-141-206-7

Query Match 39.6%; Score 38; DB 4; Length 336;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 KRIGSRLPHTGCG 15
DB 211 KRSGHRVPELTCCG 225

RESULT 15
US-08-820-170A-40
; Sequence 40, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/820,170A
;; FILING DATE:
;; CLASSIFICATION: 536
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 293-7060
;; TELEFAX: (202) 293-7860
;; TELEX: 6491103
;; INFORMATION FOR SEQ ID NO: 40:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 499 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-820-170A-40

Query Match 39.6%; Score 38; DB 2; Length 499;
Best Local Similarity 47.4%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 KRIGFS--RLPHFTGCG 17
DB 246 KRVGFSPPKDIHPMPCSL 264

Search completed: May 20, 2003, 16:44:38
Job time : 8.84615 secs

THIS PAGE BLANK (USPTO)

XX Claim 24; Page 112; 112pp; English.
XX
XX The present sequence is derived from an endo-beta-1,4-xyylanase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xyylanase protein. The xyylanase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xyylanase or
CC mutant xyylanase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xyylanase. The xyylanase inhibitor is
CC useful for screening high degree resistance xyylanases for dough
CC preparation. The xyylanase is also useful for preparing a non-sticky
CC dough. A combination of xyylanase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
XX Sequence 21 AA:
SO
Query Match 100.0%; Score 114; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 8,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LPVPAPVTKDPATSLYTIPEH 21
Db 1 LPVPAPVTKDPATSLYTIPEH 21
RESULT 2
AAE14689
ID AAE14689 standard; peptide: 21 AA.
XX
XX AAE14689;
XX
XX 21-AUG-2002 (first entry)
XX
XX Wheat flour xyylanase inhibitor B chain N-terminal fragment.
DE
XX Refrigerated dough; syruping; arabinoxylan; bakery product; bread;
KW pizza base; cake; biscuit; wheat; flour; xyylanase inhibitor.
XX
XX Triticum aestivum.
OS
XX WO200152657-A1.
PN
XX 26-JUL-2001.
PD
XX 17-JAN-2001; 2001WO-IB00168.
PF
XX 18-JAN-2000; 2000GB-0001136.
PR
XX (DANI-) DANISCO AS.
PA
XX
XX Poulsen CH, Sorensen JF;
PI
XX WPI: 2001-457446/49.
DR
XX
XX Production of refrigerated dough with reduced syruping, useful in
PT production of bakery products such as bread, comprises admixing cereal
PT flour, water and protein that prevents enzymatic degradation of
PT arabinoxylan in the cereal flour.
XX
XX Disclosure; Page 23; 26pp; English.
PS
XX The invention relates to a process for producing refrigerated dough
CC because of a reduction in water holding capacity caused by the breakdown
CC of arabinoxylan over time). The process comprises admixing cereal flour
CC and water with a protein that reduces/prevents enzymatic degradation of
CC arabinoxylan in the cereal flour. The preferred protein is a xyylanase
CC inhibitor. The method is useful to produce refrigerated dough in which
CC syruping is reduced or eliminated. Refrigerated dough is typically
CC stored for long periods to enable fresh baked products (e.g. bread,
CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
CC requirements using the method by the use of specific proteins/protein
CC combinations. The present sequence is wheat flour
CC endo-beta-1,4-xyylanase inhibitor B chain N-terminal fragment.
XX
XX Sequence 21 AA:
SO
Query Match 100.0%; Score 114; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 8,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LPVPAPVTKDPATSLYTIPEH 21
Db 1 LPVPAPVTKDPATSLYTIPEH 21
RESULT 3
AAU07394
ID AAU07394 standard; protein: 21 AA.
XX
XX AAU07394;
XX
XX 18-DEC-2001 (first entry)
XX
XX Bacillus subtilis xyylanase inhibitor #3.
DE
XX Xylanase; plant cell wall; baking; cereal; starch production; wood;
KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
XX
XX Bacillus subtilis.
OS
XX WO200166711-A1.
PN
XX 13-SEP-2001.
PD
XX 08-MAR-2001; 2001WO-IB00426.
PF
XX 08-MAR-2000; 2000GB-0005585.
PR
XX 27-JUN-2000; 2000GB-0015751.
PR
XX (DANI-) DANISCO AS.
PA
XX
XX Slbbsen O, Sorensen JF;
PI
XX WPI: 2001-596834/67.
DR
XX
XX Novel variant xyylanase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xyylanase inhibitor
XX
XX Disclosure; Page 63; 70pp; English.
PS
XX The invention relates to a variant xyylanase polypeptide (I) or its
CC fragment having xyylanase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xyylanase inhibitor as compared with the parent xyylanase enzyme. (I) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I)
CC is useful for altering the viscosity derived from the presence of
CC hemicellulose or arabinoxylan in a solution or system comprising plant
CC cell wall material. (I) is useful for preparing a foodstuff such as
CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
CC The present sequence represents the amino acid sequence of Bacillus
CC subtilis xyylanase inhibitor #3 as described in the method of the
CC invention.
XX
XX Sequence 21 AA:
SO
Query Match 100.0%; Score 114; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 8,4e-10;

XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX (LEUV-) LEUVEN RES & DEV.
 PA Delcours J, Debysers W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI: 2002-114579/15.
 FT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 FT beta-glucanolytic enzymes comprising using endoxylanases during
 FT screening for inhibition activity or affinity chromatography with
 FT immobilised enzymes -
 PS Claim 127, Page 51; 127pp; English.
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruiping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malling of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-I ('T. aestivum L. endoxylanase
 CC inhibitor').
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 XX Sequence 381 AA;
 SQ
 Query Match 91.2%; Score 104; DB 23; Length 381;
 Best Local Similarity 95.2%; Pred. No. 5.6e-07;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LPPAPVTKDPATSLYTIPIH 21
 Db 1 LPIVLAIPVTKDPATSLYTIPIH 21
 RESULT: 6
 AAU75814
 ID FAU75814 standard; Protein: 381 AA.
 XX
 AC AAU75814;
 XX

DT 23-APR-2002 (first entry)
 XX
 DE Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.
 XX
 KW Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malling;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 XX Triticum aestivum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95
 FT /Label= Unknown
 FT Misc-difference 98
 FT /Label= Unknown.
 FT Misc-difference 101
 FT /Label= Unknown
 FT Misc-difference 110
 FT /Label= Unknown
 FT Misc-difference 145
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 183
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 232
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 275
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 282
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 333
 FT /Label= Unknown
 FT
 XX WO200198474-A1.
 FN
 XX
 PD 27-DEC-2001.
 XX
 PD 21-JUN-2001; 2001WO-BE00106.
 XX
 PF 22-JUN-2000; 2000GB-0015296.
 XX
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 XX Delcours J, Debysers W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX
 DR WPI: 2002-114579/15.
 XX
 FT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 FT beta-glucanolytic enzymes comprising using endoxylanases during
 FT screening for inhibition activity or affinity chromatography with
 FT immobilised enzymes -
 PS Claim 127, Page 51; 127pp; English.
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory

CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a microheterogenic variant of wheat TAXI-I (T. aestivum L
 CC endoxylanase inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

XX SQ Sequence 381 AA;
 Query Match 91.2%; Score 104; DB 23; Length 381;
 Best Local Similarity 95.2%; Pred. No. 5.6e-07;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LPVPAPVTKDPATSLYTIPTF 21
 DB 1 LPVLAPVTKDPATSLYTIPTF 21

RESULT 7
 AAU75806 standard; peptide: 24 AA.

AC AAU75806;

DT 23-APR-2002 (first entry)

XX Wheat L endoxylanase inhibitor, TAXI I, N-terminus #1.

DE Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;

KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;

KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;

KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;

KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Triticum aestivum.

XX Key Location/Qualifiers

XX Misc-difference 21 /label= OTHER

XX Misc-difference 23 /note= "Other= Any amino acid, preferably Leu"

XX Misc-difference 23 /label= OTHER

XX /note= "Other= Any amino acid, preferably Leu"

XX WO200198474-A1.

XX 27-DEC-2001.

XX 21-JUN-2001; 2001WO-BE00106.

XX 22-JUN-2000; 2000GB-0015296.

XX 25-JAN-2001; 2001GB-0002018.

XX 16-MAR-2001; 2001GB-0005564.

XX 21-MAY-2001; 2001GB-0013238.

PA (LEUV-) LEUVEN RES & DEV.

XX Delcourt J, Delyser W, Gebruers K, Goesaert H, Fierens K, Robben J;

PI Van Campenhout S;

XX WPI: 2002-114579/15.

DR Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or

PT beta-glucanolytic enzymes comprises using endoxylanases during

PT screening for inhibitor activity or affinity chromatography with

PT immobilised enzymes

PS Claim 127; Page 9; 127pp; English.

XX The invention relates to separating and/or isolating inhibitors of

CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises

CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of

CC different specificity or by using an affinity chromatographic step with

CC immobilised enzymes and/or antibodies against inhibitors. Also

CC included are an isolated nucleic acid molecule encoding an inhibitor

CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,

CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,

CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the

CC nucleic acid, an expression system transformed with the nucleic

CC acid, a host organism transformed with the nucleic acid, the inhibitory

CC proteins encoded by the nucleic acids and modulators of the proteins.

CC A recombinant protein, glycoprotein or polypeptide or microorganisms,

CC plant or plant materials transformed with the nucleic acid are useful

CC for the formation of an endoxylanase-inhibitor complex, screening

CC endoxylanases that are totally, less or not inhibited by the inhibitors,

CC reducing syruping in refrigerated dough compositions, affecting the

CC relative affinity and/or relative hydrolysis specificity and/or relative

CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an

CC endoxylanase/inhibitor complex, improving the malting of cereals such as

CC barley, sorghum and wheat and/or the production of beer, improving the

CC production and/or quality of baked or extruded cereal products such as

CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,

CC biscuits, pasta and noodles, animal feed stuff, improving the production

CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten

CC starch separation and production, improving maize processing,

CC plant disease resistance and nutraceutical and/or pharmaceutical

CC applications, improving paper and pulp technologies. The present

CC sequence is the N-terminus of wheat TAXI-I (T. aestivum L endoxylanase

CC inhibitor) from a 40-43kDa isolate of the protein.

XX SQ Sequence 24 AA;

Query Match 84.2%; Score 96; DB 23; Length 24;

Best Local Similarity 95.0%; Pred. No. 3.9e-07;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LPVPAPVTKDPATSLYTIPTF 20
 DB 1 LPVLAPVTKDPATSLYTIPTF 20

RESULT 8
 AAU75828 standard; peptide: 22 AA.

AC AAU75828;

DT 23-APR-2002 (first entry)

XX Wheat xylanase inhibitor TAXI-I PCR product N-terminal sequence.

DE Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;

KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;

KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;

KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;

KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX Trifolium aestivum cultivar Estica.
 OS Synthetic.
 XX WO200198474-A1.
 PN
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEEVEN RES & DEV.
 FA
 FI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 FI Van Campenhout S;
 XX
 DR WPI: 2002-114579/15.
 DR N-PSDB: ABK13690.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 XX Example 11: Fig 21: 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syringing in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents the sequence surrounding the N-terminus of wheat
 CC TAXI-I (T. aestivum L endoxylanase inhibitor) which was expressed in
 CC E. coli.
 CC
 XX
 XX
 SQ Sequence 22 AA:
 Query Match 78.9%; Score 90; DB 23; Length 22;
 Best Local Similarity 94.7%; Pred. No. 2.6e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db III |||||
 4 LPLVAPVTKDPATSLYTIIP 22
 RESULT 9
 AAU75807
 ID AAU75807 standard; peptide: 22 AA.
 XX
 AC AAU75807;
 XX
 XX 23-APR-2002 (first entry)
 DT
 DE Wheat L endoxylanase inhibitor, TAXI II, N-terminus #1.
 XX
 XX
 KW Wheat; TAXI-II: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syringing; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Trifolium aestivum.
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PD 21-JUN-2001; 2001WO-BE00106.
 XX
 PF 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEEVEN RES & DEV.
 PA
 PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX
 DR WPI: 2002-114579/15.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 XX Claim 127; Page 9; 127pp; English.
 PS
 XX
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syringing in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as

CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten
CC starch separation and production, improving maize processing,
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence is the N-terminus of wheat TTXI-II ('T. aestivum L. endoxylanase
CC inhibitor) from a 40-43kDa isolate of the protein.
XX
SQ Sequence 22 AA:
Query Match 77.2%; Score 88; DB 23; Length 22;
Best Local Similarity 90.0%; Pred. No. 5.1e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 LPVPAPVTKDPATSLTYTPF 20
DB 3 LPVPAPVTKDPATSLTYTPF 22
RESULT 10
AA93764
ID AA93764 standard; peptide: 38 AA.
XX
AC AA93764;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence derived from an endo-beta-1,4-xylanase inhibitor.
XX
KW Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW dough; dough preparation.
XX
OS Triticum sp.
XX
FH Key Location/Qualifiers
FH Misc-difference 26 /note= "unspecified amino acid"
FT Misc-difference 31 /note= "unspecified amino acid"
FT Misc-difference 32 /note= "unspecified amino acid"
FT Misc-difference 38 /note= "unspecified amino acid"
XX
PN WO200039289-A2.
XX
PD 06-JUL-2000.
XX
PF 17-DEC-1999; 99WO-1B02071.
XX
PR 23-DEC-1998; 98GB-0028599.
PR 06-APR-1999; 98GB-0007805.
PR 15-APR-1999; 99GB-0008645.
XX
PA (DANI-) DANISCO AS.
XX
PI Sidsesen O, Sorensen JF;
XX
DR WPI: 2000-465744/40.
XX
PT Mutant xylanase protein identified using xylanase inhibitor useful for
PT preparing non-sticky dough for bakery products
XX
PS Disclosure: Page 106; 112pp; English.
XX
CC The present sequence is derived from an endo-beta-1,4-xylanase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xylanase protein. The xylanase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xylanase or
CC mutant xylanase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xylanase. The xylanase inhibitor is

CC useful for screening high degree resistance xylanases for dough
CC preparation. The xylanase is also useful for preparing a non-sticky
CC dough. A combination of xylanase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
SQ Sequence 38 AA:
Query Match 71.9%; Score 82; DB 21; Length 38;
Best Local Similarity 80.0%; Pred. No. 6.9e-05;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 PVPAPVTKDPATSLTYTPF 21
DB 2 PVPAPVTKDPATSLTYTPF 21
RESULT 11
AAU75808
ID AAU75808 standard; peptide: 22 AA.
XX
AC AAU75808;
XX
DT 23-APR-2002 (first entry)
XX
DE Barley L. endoxylanase inhibitor, HvXI L, N-terminus #1.
XX
KM Barley; HvXI-L; L endoxylanase inhibitor; cellulosytic enzyme inhibitor;
KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
OS Hordeum vulgare.
XX
FH Key Location/Qualifiers
FH Misc-difference 21 /label= OTHER
FT Misc-difference 22 /note= "Other- Any amino acid, preferably Pro"
FT Misc-difference 22 /label= OTHER
FT /note= "Other- Any amino acid, preferably Phe"
XX
PN WO200198474-A1.
XX
PD 27-DEC-2001.
XX
PF 21-JUN-2001; 2001WO-BE00106.
XX
PR 22-JUN-2000; 2000GB-0015296.
PR 25-JAN-2001; 2001GB-0002018.
PR 26-JAN-2001; 2001GB-0002194.
PR 16-MAR-2001; 2001GB-0006564.
PR 21-MAY-2001; 2001GB-0012328.
XX
PA (LEUV-) LEUVEN RES & DEV.
XX
PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
PI Van Campenhout S;
XX
DR WPI: 2002-114579/15.
XX
PT Separating and/or isolating inhibitors of cellulosytic, xylanolytic, or
PT beta-glucanolytic enzymes comprises using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
PS Claim 127; Page 9; 127pp; English.
XX
CC The invention relates to separating and/or isolating inhibitors of
CC cellulosytic, xylanolytic and/or beta-glucanolytic enzymes comprises
CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syrruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is the N-terminus of barley HvXI-I (L endoxylanase
 CC inhibitor) from a 40-43kDa isolate of the protein.

XX Sequence 22 AA:

Query Match 65.8%; Score 75; DB 23; Length 22;
 Best Local Similarity 88.9%; Pred. No. 0.00039;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVPVAPVTKDPATSLYTI 18
 ||| ||||| ||||| |||||
 Db 3 LPVLAPVTKDAATSLYTI 20

RESULT: 12

AAU75823 standard; Protein: 74 AA.

AC AAU75823;

DT 23-APR-2002 (first entry)

DE Maize L endoxylanase inhibitor, partial sequence ZMXI-01.

KW Maize; ZMXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syrruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX Zea mays.

OS WO200198474-A1.

PN 27-DEC-2001.

PD 21-JUN-2001; 2001WO-BE00106.

XX 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 16-MAR-2001; 2001GB-0002194.

PR 21-MAY-2001; 2001GB-0012328.

XX (LEUV-) LEUVEN RES & DEV.

XX Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX WPI: 2002-114579/15.
 DR N-PSDB: ABK13682.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 62; 127pp; English.

CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syrruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial rice ZMXI-I (Z. mays L endoxylanase
 CC inhibitor).

SO Sequence 74 AA:

Query Match 47.4%; Score 54; DB 23; Length 74;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 ATSLYTIPIPH 21
 ||||| ||||| ||||| |||||
 Db 1 ATSLYTIPIPH 10

RESULT: 13

AAU75819 standard; Protein: 287 AA.

AC AAU75819;

DT 23-APR-2002 (first entry)

XX Wheat L endoxylanase inhibitor, partial sequence TAXI-III.

DE Wheat; TAXI-III; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syrruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX Triticum aestivum cultivar Soissons.
 XX WO200198474-A1.
 XX PD 27-DEC-2001.
 XX PF 21-JUN-2001; 2001WO-BE00106.
 XX PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX (LEUV-) LEUVEN RES & DEV.
 PA Delcours J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 PI WPI: 2002-114579/15.
 DR N-PSDB: ABK13676.
 DR Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 XX beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX Claim 127; Page 58; 127pp; English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arbinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malling of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-III (T. aestivum L endoxylanase
 CC inhibitor).
 XX
 XX Sequence 287 AA:
 SO
 Query Match 47.4%; Score 54; DB 23; Length 287;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 12 ATSLYTPPH 21
 |||||
 Db 1 ATSLYTPPH 10

RESULT 14
 AAU75815
 ID AAU75815 standard; Protein: 370 AA.
 XX
 XX AAU75815;
 AC
 AC 23-APR-2002 (first entry)
 DT
 DT
 DE Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.
 XX
 XX
 XX Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malling;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 XX Triticum aestivum cultivar Soissons.
 OS
 OS WO200198474-A1.
 PN
 PN 27-DEC-2001.
 PD
 PD 21-JUN-2001; 2001WO-BE00106.
 PF
 PF 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX (LEUV-) LEUVEN RES & DEV.
 PA Delcours J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 PI WPI: 2002-114579/15.
 DR N-PSDB: ABK13672.
 DR Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 XX beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX Claim 127; Page 57; 127pp; English.
 PS
 PS The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arbinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malling of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 8.53846 Seconds
(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-15

Perfect score: 114
Sequence: 1 LPVPAVPYTKDPATSLYTRPFH 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51.5	45.2	100	T18020	hypothetical prote
2	50	43.9	3498	T22330	hypothetical prote
3	49	43.0	532	S74453	hypothetical prote
4	49	43.0	1047	A55617	masquerade precurs
5	48.5	42.5	402	F96640	hypothetical prote
6	48	42.1	273	G83049	serine acetyltrans
7	48	42.1	434	G86163	hypothetical prote
8	47	41.2	336	S75947	hypothetical prote
9	47	41.2	336	T15837	hypothetical prote
10	47	41.2	750	AB0708	hypothetical prote
11	47	41.2	1075	G59434	catalase (EC 1.11.
12	47	41.2	1199	T13946	KIAA0411 protein l
13	46	40.4	151	S10084	hypothetical 16k p
14	46	40.4	186	A70574	probable lppu prot
15	46	40.4	760	T06291	extensin homolog T
16	46	40.4	1299	T47182	hypothetical prote
17	46	40.4	1473	A5186	salivary agglutini
18	46	40.4	13055	T16580	hypothetical prote
19	45	39.5	241	T48006	hypothetical prote
20	45	39.5	275	T51437	hypothetical prote
21	45	39.5	435	A51285	glutamate-1-semial
22	45	39.5	432	E72563	probable tRNA nucl
23	45	39.5	440	JC2065	6-phosphofructo-2-
24	45	39.5	461	S45568	nuclear factor I-A
25	45	39.5	470	JC2064	6-phosphofructo-2-
26	45	39.5	498	S45567	nuclear factor I-A
27	45	39.5	509	S45565	nuclear factor I-A
28	45	39.5	522	S09996	nuclear factor I-A
29	45	39.5	597	AH2351	serine/threonine k

30	45	39.5	753	2	A39129	catalase (EC 1.11.
31	45	39.5	753	2	B85782	catalase, hydropet
32	45	39.5	753	2	F90933	catalase, HP11 limp
33	45	39.5	1048	2	T31425	C-terminal domain-
34	44.5	39.0	532	2	JT0530	muscarinic acetylch
35	44.5	39.0	532	2	T24222	hypothetical prote
36	44	38.6	187	2	S38036	hypothetical prote
37	44	38.6	357	2	PC4293	nuclear factor 1 f
38	44	38.6	390	1	Q08E77	glycoprotein I pre
39	44	38.6	416	1	A31959	lysosome-associate
40	44	38.6	426	2	T20265	hypothetical prote
41	44	38.6	460	2	T45968	hypothetical prote
42	44	38.6	504	2	A57215	glial cells missin
43	44	38.6	509	2	JC5428	nuclear factor 1 f
44	44	38.6	532	2	B36596	nuclear factor I -
45	44	38.6	549	1	NDEC	glucose-6-phosphat

ALIGNMENTS

```

RESULT 1
T18020
hypothetical protein a518r - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18020
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18020
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-100 <GRA>
A:Cross-References: EMBL:U42580; NID:94028896; PIDN:AAC36885.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: a518r

Query Match      45.2%; Score 51.5; DB 2; Length 100;
Best Local Similarity 45.8%; Pred. No. 1.2;
Matches 11; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

OY      1 LPVPA---PYTKDPATSLYTRPFH 21
DB      17 VPPPAVEKPYQGHPAVWMTLPFH 40

RESULT 2
T22330
hypothetical protein F47A4.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T22330
R:Mortimore, B.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19549
A:Accession: T22330
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-3498 <WIL>
A:Cross-References: EMBL:U49888; PIDN:CAA90064.1; GSPDB:GN00028; CESP:F47A4.2
A:Experimental source: clone F47A4
C:Genetics:
A:Gene: CESP:F47A4.2
A:Map position: X
A:Introns: 36/3; 72/3; 147/2; 228/3; 261/2; 320/2; 847/3; 881/1; 1317/2; 1500/3; 1653

Query Match      43.9%; Score 50; DB 2; Length 3498;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      2 PVPAPVTKDPAT 13

```

Db 2305 PKPAPYTKDPATSPAR 2316

RESULT 3

hypothetical protein slr1484 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S74453

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shilpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

5

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74453

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-532 <KAN>

A:Cross-references: EMBL:D90899; GB:AB001339; NID:91651650; PIDN:BAAL6605.1; PID:dl01733

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 43.0%; Score 49; DB 2; Length 532;

Best Local Similarity 40.0%; Pred. No. 20;

Matches 10; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

OY 2 PVPAPYTKDPATSPAR-----SLYTIPTF 20

Db 161 PLNPVTESPAITOPAQSAFLRF 185

RESULT 4

masquerade precursor - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 21-Jul-2000

C:Accession: A55617

R:Murugasu-Oel, B.; Rodrigues, V.; Yang, X.; Chia, W.

Genes Dev. 9, 139-154, 1995

A:Title: Masquerade: a novel secreted serine protease-like molecule is required for some

A:Reference number: A55617; MUID:95154720; PMID:7851790

A:Accession: A55617

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1047 <RES>

A:Cross-references: EMBL:U18130; NID:g665544; PIDN:ANCA6512.1; PID:g665545

C:Comment: This secreted protein, unlikely to function as a serine proteinase, is proces

A:Genes: mas

A:Cross-references: FlyBase:FBgn0011653

C:Superfamily: trypsin homology

F:803-1038/Domain: trypsin homology <TRY>

Query Match 43.0%; Score 49; DB 2; Length 1047;

Best Local Similarity 61.1%; Pred. No. 43;

Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 PVPAPYTKDPATSPATSLYTIPTF 19

Db 141 PKKPOTKRPATSPSTTKP 158

RESULT 5

hypothetical protein T25B24.11 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: F96640

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A66141; MUID:21016719; PMID:11130712

A:Accession: F96640

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-402 <STO>

A:Cross-references: GB:AE005173; NID:94585881; PIDN:AAD25554.1; GSPDB:GN00141

C:Genetics:

A:Gene: T25B24.11

A:Map position: 1

C:Superfamily: caffeoyl-CoA 3-O-methyltransferase

Query Match 42.5%; Score 48.5; DB 2; Length 402;

Best Local Similarity 40.0%; Pred. No. 17;

Matches 8; Conservative 8; Mismatches 3; Indels 1; Gaps 1;

OY 1 LPVPAPYTKDPATSPATSLYTIPTF 20

Db 14 MPYSPVTEBP-ISTMSLPY 32

RESULT 6

serine acetyltransferase VC2649 [imported] - *Vibrio cholerae* (strain N16961 serogroup

C:Species: *Vibrio cholerae*

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: G82049

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: G82049

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-273 <HEI>

A:Cross-references: GB:AE004331; GB:AE003852; NID:g9657236; PIDN:AAFP5790.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2649

A:Map position: 1

C:Superfamily: serine acetyltransferase; serine acetyltransferase homology

Query Match 42.1%; Score 48; DB 2; Length 273;

Best Local Similarity 53.3%; Pred. No. 13;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 APYTKDPATSPATSLYTIPTF 19

Db 87 ATVNRDPAVMSWSP 101

RESULT 7

hypothetical protein F15K9.16 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: G86163

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marcia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86163
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 <STO>
 A:Cross-References: GB:AE005172; NID:g3850580; PIDN:AACT2120.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: conglutin gamma

Query Match 42.1%; Score 48; DB 2; Length 434;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 PVTKDPATSLYTI 17
 |||||:|
 DB 35 PVTKDPSTLOYT 46

RESULT 8
 S75947
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: S75947
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 S.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S75947
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-336 <KAN>
 A:Cross-References: EMBL:D64006; GB:AB001339; NID:g1001291; PID:g1001307
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: conserved hypothetical protein ylnf

Query Match 41.2%; Score 47; DB 1; Length 336;
 Best Local Similarity 47.4%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 VPAVPVKDPATSLYTIPT 21
 |||||:|
 DB 6 VPAVPSLFPETLEPLPYH 24

RESULT 9
 T15837
 hypothetical protein C54D2.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15837
 R:Minx, P.
 submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of C. elegans cosmid C54D2.
 A:Reference number: Z18415
 A:Accession: T15837
 A:Status: preliminary; translated from GB/EMBL/DDJ
 A:Molecule type: DNA
 A:Residues: 1-336 <MIN>
 A:Cross-References: EMBL:U37548; NID:g1017804; PID:g1017808; PIDN:AAAY9200.1; CESP:C54D2
 C:Genetics:
 A:Gene: CESP:C54D2.1
 A:Introns: 31/1; 87/1; 116/1; 150/2; 221/3; 295/3

Query Match 41.2%; Score 47; DB 2; Length 336;
 Best Local Similarity 52.6%; Pred. No. 23;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 LPVPAPVTKDPATSLYTI 19
 :|||:|
 DB 116 VPVPVPTATPATVAPTPAP 134

RESULT 10
 AB0708
 catalase (EC 1.11.1.6) - *Salmonella enterica* subsp. *enterica* serovar Typh1 (strain CT
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh1
 A:Note: This species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 17-May-2002
 C:Accession: AB0708
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 , S.; Mout, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AB0708
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-750 <PAR>
 A:Cross-References: GB:AL513382; PIDN:CAD02035.1; PID:g16502872; GSPDB:GN00176
 C:Genetics:
 A:Gene: SRY1793
 C:Superfamily: catalase
 C:Keywords: oxidoreductase

Query Match 41.2%; Score 47; DB 2; Length 750;
 Best Local Similarity 50.0%; Pred. No. 58;
 Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

OY 4 PAP---VTKDPATSLYTI 19
 | : |||||
 DB 572 PPDPVNGLRKDPALSLYAVP 591

RESULT 11
 G59434
 KIAA0411 protein [imported] - human
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
 C:Accession: G59434
 R:Turner, A.C.; Zambrowicz, B.; Nehls, M.; Friedrich, G.A.; Sands, A.T.
 patent application, Lexicon Genetics Incorporated (US), WO 0075320-A, December 2000
 A:Description: Human genes and proteins encoded thereby.
 A:Reference number: G59434
 A:Accession: G59434
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1075 <TUR>
 A:Cross-References: GB:CAZ22407; PID:g12309846; PIDN:CAZ22407.1

Query Match 41.2%; Score 47; DB 2; Length 1075;
 Best Local Similarity 53.3%; Pred. No. 89;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 4 PAPVTKDPATSLYTI 18
 ||||:|
 DB 978 PGVPSSEPSAPLHTI 992

RESULT 12
 T13946
 probable adaptor-related protein complex AP-3, delta chain - mouse
 N:Alternate names: bovine leukemia virus receptor homolog
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13946
 R:Suzuki, T.; Ikeda, H.
 J. Virol. 72, 593-599, 1998

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 5.65385 Seconds

(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-15

Perfect score: 114

Sequence: 1.LPVPAVTKDPATSLYTRPFH 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	43.9	349	1	IRF2_MOUSE
2	48	42.1	179	1	RBS1_FRIG
3	48	42.1	179	1	RBS3_FRIG
4	47	41.2	650	1	Y411_HUMAN
5	46	40.4	151	1	WTIS_WHEAT
6	46	40.4	555	1	DABL1_HUMAN
7	46	40.4	555	1	DABL1_MACR
8	46	40.4	806	1	MK07_MOUSE
9	45	39.5	328	1	HXCD_MOUSE
10	45	39.5	330	1	HXCD_MOUSE
11	45	39.5	470	1	P26_RANCA
12	45	39.5	509	1	NR1A_HUMAN
13	45	39.5	522	1	NR1A_CHICK
14	45	39.5	753	1	CATE_ECOLI
15	45	39.5	1048	1	SRA4_RAT
16	44.5	39.0	532	1	ACM5_HUMAN
17	44	38.6	187	1	YKT9_YEAST
18	44	38.6	390	1	VGLI1_HSV1
19	44	38.6	416	1	LMP1_HUMAN
20	44	38.6	509	1	NR1A_RAT
21	44	38.6	532	1	NR1A_MOUSE
22	44	38.6	549	1	G6P1_ECOLI
23	44	38.6	562	1	LCB2_KLUL
24	44	38.6	669	1	C101_RAT
25	44	38.6	1844	1	POLR_TYRV
26	44	38.6	1844	1	POLR_TYRV
27	44	38.6	1844	1	POLR_TYRV
28	43.5	38.2	878	1	IL3B_MOUSE
29	43	37.7	116	1	UBI1_HICMA
30	43	37.7	235	1	MYF5_XENLA
31	43	37.7	402	1	MYF5_XENLA
32	43	37.7	413	1	CRTX_ERME
33	43	37.7	582	1	MNT_HUMAN

34	43	37.7	1157	1	SRA4_HUMAN	O95104 homo sapien
35	43	37.7	1253	1	SHK2_HUMAN	O949X8 homo sapien
36	43	37.7	1324	1	POL2_GCMV	P13026 hungarian g
37	43	37.7	1474	1	SHK2_RAT	O94974 rattus norv
38	43	37.7	3591	1	FNAB_BOPE	P12255 bordelella
39	42.5	37.3	215	1	HIS5_DEIRA	O91X89 deinococcus
40	42.5	37.3	432	1	ST11_XENLA	O91604 xenopus lae
41	42.5	37.3	601	1	DNM1_MYGE	P47442 mycoplasma
42	42.5	37.3	906	1	CB1_HUMAN	P22681 homo sapien
43	42	36.8	182	1	YR31_CAEBL	O10958 caenorhabdi
44	42	36.8	273	1	ROCI1_NICSY	O08935 nicotiana s
45	42	36.8	401	1	NNOS_DROME	P25724 drosophila

ALIGNMENTS

RESULT 1
ID IRF2_MOUSE STANDARD: PRT: 349 AA.
AC P23906:
DF 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DR 30-MAY-2000 (Rel. 39, Last annotation update)
DE Interferon regulatory factor 2 (IRF-2).
GN IRF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89354547; PubMed=2475256;
RA Harada H., Fujita T., Miyamoto M., Kimura Y., Maruyama M.,
RA Furia A., Miyata T., Taniguchi T.;
RT "Structurally similar but functionally distinct factors, IRF-1 and
RT IRF-2, bind to the same regulatory elements of IFN and IFN-inducible
RT genes";
RT Cell 58:729-739(1989).
RN [2]
RP STRUCTURE BY NMR OF 2-113.
RX MEDLINE=98230747; PubMed=9562558;
RA Furi J., Uegaki K., Yamazaki T., Shirakawa M., Swindells M.B.,
RA Harada H., Taniguchi T., Kyogoku Y.;
RT "Solution structure of the IRF-2 DNA-binding domain: a novel subgroup
RT of the winged helix-turn-helix family";
RT Structure 6:491-500(1998).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-113.
RX MEDLINE=99417584; PubMed=10487755;
RA Fujii Y., Shimizu T., Kusumoto M., Kyogoku Y., Taniguchi T.,
RA Hakoshima T.;
RT "Crystal structure of an IRF-DNA complex reveals novel DNA recognition
RT and cooperative binding to a tandem repeat of core sequences";
RT EMBO J. 18:5028-5041(1999).
RL
CC -!- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION OF
CC TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON
CC CONSENSUS SEQUENCE (ICS1)) AND REPRESSES THOSE GENES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: BY VIRUSES AND IFN.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J03168; AAA39333.1;
CC PIR: A32828; A32828.
CC DR
PDB: 1IRF; 28-JAN-96.

DR DB: 21RF; 08-OCT-99.
 DR DB: 11RG; 18-MAR-98.
 DR TRANSFAC; T00425; -.
 DR MGD; MGI:96591; 1rf2.
 DR InterPro: IPR001346; 1RF.
 DR Pfam: PF00605; 1RF; 1.
 DR PRINTS; PR00267; INTERPRECT.
 DR ProDom; PD002355; 1RF; 1.
 DR SMART; SM00348; 1RF; 1.
 DR PROSITE; PS00601; 1RF; 1.
 KW transcription regulation; DNA-binding; Repressor; Nuclear protein;
 KW Interferon induction; 3D-structure.
 FT DNA_BIND 7 109 TRYPTOPHAN PENTAD REPEAT.
 SQ SEQUENCE 349 AA; 39452 MW; 8738B082FB40FB11 CRC64;

Query Match 43.9%; Score 50; DB 1; Length 349;
 Best Local Similarity 64.3%; Pred. NO. 5; Mismatches 2; Indels 0; Gaps 0;
 Matches 9; Conservative 3;

OY 1 LPVPAPVTKDPATS 14
 11:11111 1:1
 Db 311 LPVPAPVTKDPSSS 324

RESULT 2
 RBS1_FRIAG STANDARD; PRT; 179 AA.
 AC 024634:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribulose biphosphate carboxylase small chain 1/4, chloroplast
 DE precursor (EC 4.1.1.39) (Rubisco small subunit 1/4).
 GN RBCS4.
 OS Fritillaria agrestis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Fritillaria.
 OC NCBI_TaxID=64177;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Panico E., Baysdorfer C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 ACTIVE SITE.
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) -> 2 3-
 phospho-D-glycerate.
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) ->
 3-phospho-D-glycerate + 2-phosphoglycolate.
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF031543; AAB86853.1; -.
 CC DR EMBL; AF024572; AAB84179.1; -.
 CC DR InterPro; IPR000894; Rubisco_small.
 CC DR Pfam; PF00101; Rubisco_small; 1.
 CC DR PRINTS; PR00152; RUBISCO_SMALL.
 CC DR ProDom; PD000290; Rubisco_small; 1.
 CC DR ProCosynthesis; Carbon dioxide fixation; Photorespiration;
 KW Phycosynthesis; Carbon dioxide fixation; Photorespiration;
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transist peptide;
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transist peptide;

KW MultiGene family.
 FT TRANSIT 1 59
 FT CHAIN 60 179
 FT
 SQ SEQUENCE 179 AA; 19620 MW; 1695775BD611A765 CRC64;

Query Match 42.1%; Score 48; DB 1; Length 179;
 Best Local Similarity 64.3%; Pred. NO. 4.8; Mismatches 3; Indels 0; Gaps 0;
 Matches 9; Conservative 2;

OY 6 PVTGKPAATSLYTP 19
 111:11111 1:1
 Db 38 PVTGKPAATSLTP 51

RESULT 3
 RBS3_FRIAG STANDARD; PRT; 179 AA.
 AC 022573:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribulose biphosphate carboxylase small chain 3, chloroplast precursor
 DE (EC 4.1.1.39) (Rubisco small subunit 3).
 GN RBCS3.
 OS Fritillaria agrestis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Fritillaria.
 OC NCBI_TaxID=64177;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Panico E., Baysdorfer C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 ACTIVE SITE.
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) -> 2 3-
 phospho-D-glycerate.
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) ->
 3-phospho-D-glycerate + 2-phosphoglycolate.
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF024574; AAB84181.1; -.
 CC DR InterPro; IPR000894; Rubisco_small.
 CC DR Pfam; PF00101; Rubisco_small; 1.
 CC DR PRINTS; PR00152; RUBISCO_SMALL.
 CC DR ProDom; PD000290; Rubisco_small; 1.
 CC DR ProCosynthesis; Carbon dioxide fixation; Photorespiration;
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transist peptide;
 KW MultiGene family.
 KW TRANSIT 1 59
 KW CHAIN 60 179
 KW
 FT CHAIN 60 179
 FT
 SQ SEQUENCE 179 AA; 19615 MW; D2C68A6CC0E08F66 CRC64;

Query Match 42.1%; Score 48; DB 1; Length 179;
 Best Local Similarity 64.3%; Pred. NO. 4.8; Mismatches 3; Indels 0; Gaps 0;
 Matches 9; Conservative 2;

Oy 6 PVTKDPATSLYTP 19
 Db 38 PVTOKPATGLSTLP 51

RESULT 4

Y411_HUMAN STANDARD; PRT; 650 AA.
 AC 043295;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein KIAA0411 (Fragment).
 GN KIAA0411.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9811655; PubMed=9455477;
 RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VIII.
 RT 78 new cDNA clones from brain which code for large proteins in
 RT vitro.";
 RL DNA Res. 4:307-313(1997).
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB007871; BAA24841.1; -
 DR HSSP: O60631; IGBO.
 DR InterPro: IPR000198; RhoGAP.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhoGAP; 1.
 DR ProDom: PD000066; SH3; 1.
 DR SMART: SM00324; RhoGAP; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR Hypothetical protein; GTPase activation; SH3 domain.
 FT NON_TER 1 1
 FT DOMAIN 71 214 RHO-GAP.
 FT DOMAIN 296 365 SH3.
 SO SEQUENCE 650 AA; 71863 MW; 7175AA9CCF259907 CRC64;
 Query Match 41.2%; Score 47; DB 1; Length 650;
 Best Local Similarity 53.3%; Pred. No. 26;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 4 PAPVTKDPATSLYTP 18
 Db 553 PGVPSSEBASPLHTI 567

RESULT 5

Y411_HUMAN STANDARD; PRT; 151 AA.
 AC 043295;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 16 kDa protein in middle repetitive insertion sequence
 DE WIS1.

OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 CC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Chinese Spring;
 RX MEDLINE=89364725; PubMed=2549380;
 RA Martienssen R.A., Paulcombe D.C.;
 RT "An unusual wheat insertion sequence (WIS1) lies upstream of an
 RT alpha-amylase gene in hexaploid wheat, and carries a 'minisatellite'
 RT array.";
 RL Mol. Gen. Genet. 217:401-410(1989).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X15870; CAA33880.1; -
 DR PIR: S10084; S10084.
 KW Transposable element; Hypothetical protein.
 SO SEQUENCE 151 AA; 15950 MW; 9BAB30BD31EC6742 CRC64;

Query Match 40.4%; Score 46; DB 1; Length 151;
 Best Local Similarity 47.4%; Pred. No. 7.7;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Oy 2 PVPATKDPATSLYTP 20
 Db 64 PPPPTSPSISLPPF 82

RESULT 6

DAB1_HUMAN STANDARD; PRT; 555 AA.
 AC 073553; Q9NYA8;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Disabled homolog 1.
 GN DAB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99009343; PubMed=9790777;
 RA Lambert de Rouvroit C., Goffinet A.M.;
 RT "Cloning of human DAB1 and mapping to chromosome 1p31-p32."
 RT Genomics 53:246-247(1998).
 RL [2]
 SO SEQUENCE FROM N.A.
 RP Fazili Z., Sun W., Xu X.-X.;
 RA "Aberrant disabled-1 expression in tumors."
 RT Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Adaptor molecule functioning in neural development (By
 CC similarity).
 CC -1- SUBUNIT: Associates with the SH2 domains of Src, Fyn and Abl (By
 CC similarity).
 CC -1- DOMAIN: THE PID DOMAIN SPECIFICALLY BINDS TO THE ASN-PRO-XAA-
 CC TYR(P) MOTIF FOUND IN MANY TYROSINE-PHOSPHORYLATED PROTEINS.
 CC -1- PTM: Phosphorylated on tyrosine (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

KW Phosphorylation. 347 PROTEIN KINASE.
FT DOMAIN 55 465 PRO-RICH 1.
FT DOMAIN 434 465 POLY-ARG.
FT DOMAIN 521 524
FT DOMAIN 578 700 PRO-RICH 2.
FT NP_BIND 61 69 ATP (BY SIMILARITY).
FT BINDING 84 84 ATP (BY SIMILARITY).
FT ACT_SITE 182 182 BY SIMILARITY.
FT MOD_RES 219 219 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 221 221 (BY SIMILARITY).
FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 221 221 (BY SIMILARITY).
SQ SEQUENCE 806 AA: 87732 MW: E7CC41C4BDE0633 CRC64;
Query Match 40.4%; Score 46; DB 1; Length 806;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 2 PVPAPVTKDPATSL 15
DB 640 PIPAPLQTAFTSL 653
RESULT 9
HXCD_MOUSE STANDARD: PRT: 328 AA.
ID HXCD_MOUSE 092017;
AC P50207; 092017;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Homeobox protein Hox-C13.
GN HOXC13 OR HOXC-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FVB;
RA Tkatchenko A.V., Visconti R.P., Shang L., Papenbrock T., Ito T.,
RA Ogawa M., Awgulewitsch A.;
RT "Hyperproliferation of keratinocytes, alopecia and ichthyosis-like
RT syndrome in transgenic mice overexpressing Hoxc13.";
RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 272-309 FROM N.A.
RC STRAIN-C57BL/6J;
RX MEDLINE=95048380; PubMed=7959778;
RA Birdsaw M.S., Ruddle F.H.;
RT "Identification of the murine Hox-c12 and Hox-c13 homeoboxes on yeast
RT artificial chromosomes.";
RL Genomics 22:234-236(1994).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEOBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF193796; AAL09298.1; -
DR EMBL: U04838; AAL20229.1; -
DR TRNSPAC: T03354; -
DR MGI: 99560; Hoxc13.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 2.
DR ProDom: PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS00071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA_BIND 258 317 HOMEOBOX.
FT DOMAIN 27 60 GLY-RICH.
SQ SEQUENCE 328 AA: 35192 MW: AE363563B5AEF776 CRC64;
Query Match 39.5%; Score 45; DB 1; Length 328;
Best Local Similarity 52.6%; Pred. No. 25;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OY 2 PVPAPVTKDPATSLYIPF 20
DB 105 PPAPVTSSTATLGYIPF 123
RESULT 10
HXCD_HUMAN STANDARD: PRT: 330 AA.
ID HXCD_HUMAN 09NRYD5; 09NR24; 096032;
AC P31276; 09NRYD5; 09NR24; 096032;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Homeobox protein Hox-C13 (Hox-3G).
GN HOXC13 OR HOXC3G.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296799; PubMed=10835276;
RA de Stanchina E., Gabellini D., Norio P., Giacca M., Peverelli F.A.,
RA Riva S., Falaschi A., Diamonti G.;
RT "Selection of homeotic proteins for binding to a human DNA replication
RT origin.";
RL J. Mol. Biol. 299:667-680(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Li S., Zhang B., Li X.;
RT "A homeobox protein interacting with promoter region of p21WAF1/Cip1
RT gene is identical to Hox-C13 (Hox 3G).";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strussberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 260-325 FROM N.A.
RX MEDLINE=90098876; PubMed=2574852;
RA Acampora D., D'Esposito M., Faiella A., Pannese M., Migliacchio E.,
RA Morelli F., Stornaiuolo A., Nigro V., Simeone A., Boncinelli E.;
RT "The human HOX gene family.";
RL Nucleic Acids Res. 17:10385-10402(1989).
RN [5]
RP SEQUENCE OF 185-330 FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagahtsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEOBOX FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF255676; AAF67760.1; -
DR EMBL: AF263466; AAF73439.1; -
DR EMBL: BC002754; AAH02754.1; -
DR EMBL: AK024027; BAB14786.1; -
DR PIR: S14934; S14934.
DR HSSP: P14653; 1B72.
DR TRNSPAC: T03330; -
DR Genew: HGNC:5125; HOKC13.
DR MIM: 142976; -
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PSS00027; HOMEBOX_1; 1.
DR PROSITE: PSS00071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
DR Transcription regulation; Polymorphism.
FT DNA_BIND 260 319 HOMEBOX.
FT DOMAIN 27 61 GLY-RICH.
FT VARIANT 50 50 S->I (IN DBSNP:1867298).
FT CONFLICT 95 95 D->E (IN REF.1).
FT SEQUENCE 330 AA; 35379 MW; 7D67C5F1E5E4E915 CMC64;

Query Match 39.5%; Score 45; DB 1; Length 330;
Best Local Similarity 52.6%; Pred. No. 25;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTIPE 20
   1 1111 11 11
Db 107 PPPAPPTSSSATLGGYGF 125

RESULT 11
F26_RANCA STANDARD: PRT; 470 AA.
ID AC Q91309; Q91310;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6PF-2-K/Fru-2,6-P2ASE liver/muscle Isozymes [Includes: 6-
DE phosphofructo-2-kinase (EC 2.7.1.105); Fructose-2,6-bisphosphatase
DE (EC 3.1.3.46)].
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver, and muscle;
RX MEDLINE=94161720; PubMed=7509597;
RA Sakai A., Matanabe F., Furuya E.;
RT "Cloning of cDNAs for fructose 6-phosphate 2-kinase/fructose 2,6-
RT biphosphatase from frog skeletal muscle and liver, and their
RT expression in skeletal muscle.";
RL Biochem. Biophys. Res. Commun. 198;1099-1106(1994).
CC -1- FUNCTION: SYNTHESIS AND DEGRADATION OF FRUCTOSE 2,6-BISPHOSPHATE.
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 2,6-bisphosphate.
CC -1- CATALYTIC ACTIVITY: D-fructose 2,6-bisphosphate + H(2)O = D-
CC fructose 6-phosphate + phosphate.
CC -1- ENZYME REGULATION: PHOSPHORYLATION RESULTS IN INHIBITION OF THE
CC KINASE ACTIVITY (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

```

```

CC -----
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; LIVER/L-TYPE (SHOWN HERE) AND
CC MUSCLE/M-TYPE; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE
CC PHOSPHOGLYCERATE MUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D25223; BAA04952.1; -
DR EMBL: D25222; BAA04951.1; -
DR HSSP: P07953; ITIP.
DR InterPro: IPR000546; 6PF2K.
DR InterPro: IPR003094; 6PF2K_kin.
DR InterPro: IPR001345; PG/BPGM_mutase.
DR Pfam: PF00300; PGAM; 1.
DR Pfam: PF01591; 6PF2K; 1.
DR PRINTS: PR00991; 6PF2K.
DR PRODOM: PD002665; 6PF2K; 1.
DR PROSITE: PS00175; PG_MUTASE; 1.
DR PROSITE: PS00175; PG_MUTASE; 1.
DR Multifunctional enzyme; Transferase; Kinase; Hydrolyase; ATP-binding;
DR Phosphorylation; Liver; Alternative splicing.
FT DOMAIN 1 249 6-PHOSPHOFRUCTO-2-KINASE.
FT DOMAIN 250 470 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
FT MOD_RES 31 31 ATP (POTENTIAL).
FT NP_BIND 47 54 TO FRU-6-P (BY SIMILARITY).
FT BINDING 104 104 POTENTIAL.
FT ACT_SITE 130 130 POTENTIAL.
FT ACT_SITE 160 160 POTENTIAL.
FT BINDING 195 195 TO FRU-6-P (BY SIMILARITY).
FT ACT_SITE 258 258 FORMS THE PHOSPHOHISTIDINE INTERMEDIATE.
FT ACT_SITE 327 327 POTENTIAL.
FT ACT_SITE 392 392 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 392 392 MADRLRELTOTRLQKIMVPHQCDRLQQRGS -> MEGNRYK
FT VARSPLIC 1 31 LLEDKASRIPA (IN MUSCLE ISOFORM).
FT SEQUENCE 470 AA; 54932 MW; 53112EF3B86BD310 CMC64;

Query Match 39.5%; Score 45; DB 1; Length 470;
Best Local Similarity 40.0%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTIPEH 21
   1 1111 11 11
Db 450 PLNVEYSRDPEDALDTVPEH 469

RESULT 12
NFTA_HUMAN STANDARD: PRT; 509 AA.
ID AC Q12857; Q9P2A9; Q9H3X9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear factor 1 A-type (Nuclear factor 1/A) (NFI-A) (NF-1/A)
DE (CCAT-box binding transcription factor) (CTF) (TGCCA-binding
DE protein).
OS Homo sapiens (Human).
GN NFI1 OR KIA1439.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";

```


RX MEDLINE-97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.R., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12";
 RL Science 277:1453-1474(1997).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itch T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takada J.,
 RA Takeuchi K., Takeuchi Y., Mada C., Yamamoto Y., Horinouchi T.,
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map";
 RL DNA Res. 3:363-377(1996).
 [4]
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE-95393020; PubMed=7663946;
 RA Bravo J., Verdaguer N., Tormo J., Betzel C., Switala J., Loewen P.C.,
 RA Fita I.,
 RT "Crystal structure of catalase HPII from *Escherichia coli*";
 RL Structure 3:491-502(1995).
 [5]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE-99190072; PubMed=10091651;
 RA Sevinc M.S., Mate M.J., Switala J., Fita I., Loewen P.C.,
 RT "Role of the lateral channel in catalase HPII of *Escherichia coli*";
 RL Protein Sci. 8:490-498(1999).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (1.89 ANGSTROMS).
 RX MEDLINE-21348730; PubMed=11455600.
 RA Melik-Adamyan W.R., Bravo J., Carpene X., Switala J.,
 RA Mate M.J., Fita I., Loewen P.C.,
 RT "Substrate flow in catalases deduced from the crystal structures of
 active site variants of HPII from *Escherichia coli*";
 RL Proteins 44:270-281(2001).
 CC -I- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
 TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
 CC -I- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
 CC -I- COFACTOR: HEME GROUP.
 CC -I- SUBUNIT: HOMOTETRAMER.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -I- INDUCTION: BY ENTRY INTO STATIONARY PHASE.
 CC -I- SIMILARITY: BELONGS TO THE CATALASE FAMILY. HPII SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M55161; AAA24039.1; -
 DR EMBL; AE000266; AAC74802.1; -
 DR EMBL; D90815; BAA20916.1; -
 DR EMBL; D90816; BAA15513.1; -
 DR EMBL; D90817; BAA15521.1; -
 DR PIR; A39129; A39129.
 DR PDB; 1JPH; 04-SEP-97.
 DR PDB; 1CF9; 06-APR-99.
 DR PDB; 1OF7; 26-APR-99.
 DR PDB; 1G69; 30-AUG-00.
 DR PDB; 1GGE; 30-AUG-00.
 DR PDB; 1GGF; 30-AUG-00.
 DR PDB; 1GGH; 30-AUG-00.
 DR PDB; 1GGJ; 30-AUG-00.
 DR PDB; 1GGK; 30-AUG-00.

DR EcGene: EG10509; kate.
 DR InterPro: IPR002226; Catalase.
 DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PR00067; CATALASE.
 DR ProDom: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE_1; 1.
 DR PROSITE: PS00438; CATALASE_2; 1.
 DR PROSITE: PS00439; CATALASE_3; 1.
 DR PROSITE: PS00440; CATALASE_4; 1.
 DR PROSITE: PS00441; CATALASE_5; 1.
 DR PROSITE: PS00442; CATALASE_6; 1.
 DR PROSITE: PS00443; CATALASE_7; 1.
 DR PROSITE: PS00444; CATALASE_8; 1.
 DR PROSITE: PS00445; CATALASE_9; 1.
 DR PROSITE: PS00446; CATALASE_10; 1.
 DR PROSITE: PS00447; CATALASE_11; 1.
 DR PROSITE: PS00448; CATALASE_12; 1.
 DR PROSITE: PS00449; CATALASE_13; 1.
 DR PROSITE: PS00450; CATALASE_14; 1.
 DR PROSITE: PS00451; CATALASE_15; 1.
 DR PROSITE: PS00452; CATALASE_16; 1.
 DR PROSITE: PS00453; CATALASE_17; 1.
 DR PROSITE: PS00454; CATALASE_18; 1.
 DR PROSITE: PS00455; CATALASE_19; 1.
 DR PROSITE: PS00456; CATALASE_20; 1.
 DR PROSITE: PS00457; CATALASE_21; 1.
 DR PROSITE: PS00458; CATALASE_22; 1.
 DR PROSITE: PS00459; CATALASE_23; 1.
 DR PROSITE: PS00460; CATALASE_24; 1.
 DR PROSITE: PS00461; CATALASE_25; 1.
 DR PROSITE: PS00462; CATALASE_26; 1.
 DR PROSITE: PS00463; CATALASE_27; 1.
 DR PROSITE: PS00464; CATALASE_28; 1.
 DR PROSITE: PS00465; CATALASE_29; 1.
 DR PROSITE: PS00466; CATALASE_30; 1.
 DR PROSITE: PS00467; CATALASE_31; 1.
 DR PROSITE: PS00468; CATALASE_32; 1.
 DR PROSITE: PS00469; CATALASE_33; 1.
 DR PROSITE: PS00470; CATALASE_34; 1.
 DR PROSITE: PS00471; CATALASE_35; 1.
 DR PROSITE: PS00472; CATALASE_36; 1.
 DR PROSITE: PS00473; CATALASE_37; 1.
 DR PROSITE: PS00474; CATALASE_38; 1.
 DR PROSITE: PS00475; CATALASE_39; 1.
 DR PROSITE: PS00476; CATALASE_40; 1.
 DR PROSITE: PS00477; CATALASE_41; 1.
 DR PROSITE: PS00478; CATALASE_42; 1.
 DR PROSITE: PS00479; CATALASE_43; 1.
 DR PROSITE: PS00480; CATALASE_44; 1.
 DR PROSITE: PS00481; CATALASE_45; 1.
 DR PROSITE: PS00482; CATALASE_46; 1.
 DR PROSITE: PS00483; CATALASE_47; 1.
 DR PROSITE: PS00484; CATALASE_48; 1.
 DR PROSITE: PS00485; CATALASE_49; 1.
 DR PROSITE: PS00486; CATALASE_50; 1.
 DR PROSITE: PS00487; CATALASE_51; 1.
 DR PROSITE: PS00488; CATALASE_52; 1.
 DR PROSITE: PS00489; CATALASE_53; 1.
 DR PROSITE: PS00490; CATALASE_54; 1.
 DR PROSITE: PS00491; CATALASE_55; 1.
 DR PROSITE: PS00492; CATALASE_56; 1.
 DR PROSITE: PS00493; CATALASE_57; 1.
 DR PROSITE: PS00494; CATALASE_58; 1.
 DR PROSITE: PS00495; CATALASE_59; 1.
 DR PROSITE: PS00496; CATALASE_60; 1.
 DR PROSITE: PS00497; CATALASE_61; 1.
 DR PROSITE: PS00498; CATALASE_62; 1.
 DR PROSITE: PS00499; CATALASE_63; 1.
 DR PROSITE: PS00500; CATALASE_64; 1.
 DR PROSITE: PS00501; CATALASE_65; 1.
 DR PROSITE: PS00502; CATALASE_66; 1.
 DR PROSITE: PS00503; CATALASE_67; 1.
 DR PROSITE: PS00504; CATALASE_68; 1.
 DR PROSITE: PS00505; CATALASE_69; 1.
 DR PROSITE: PS00506; CATALASE_70; 1.
 DR PROSITE: PS00507; CATALASE_71; 1.
 DR PROSITE: PS00508; CATALASE_72; 1.
 DR PROSITE: PS00509; CATALASE_73; 1.
 DR PROSITE: PS00510; CATALASE_74; 1.
 DR PROSITE: PS00511; CATALASE_75; 1.
 DR PROSITE: PS00512; CATALASE_76; 1.
 DR PROSITE: PS00513; CATALASE_77; 1.
 DR PROSITE: PS00514; CATALASE_78; 1.
 DR PROSITE: PS00515; CATALASE_79; 1.
 DR PROSITE: PS00516; CATALASE_80; 1.
 DR PROSITE: PS00517; CATALASE_81; 1.
 DR PROSITE: PS00518; CATALASE_82; 1.
 DR PROSITE: PS00519; CATALASE_83; 1.
 DR PROSITE: PS00520; CATALASE_84; 1.
 DR PROSITE: PS00521; CATALASE_85; 1.
 DR PROSITE: PS00522; CATALASE_86; 1.
 DR PROSITE: PS00523; CATALASE_87; 1.
 DR PROSITE: PS00524; CATALASE_88; 1.
 DR PROSITE: PS00525; CATALASE_89; 1.
 DR PROSITE: PS00526; CATALASE_90; 1.
 DR PROSITE: PS00527; CATALASE_91; 1.
 DR PROSITE: PS00528; CATALASE_92; 1.
 DR PROSITE: PS00529; CATALASE_93; 1.
 DR PROSITE: PS00530; CATALASE_94; 1.
 DR PROSITE: PS00531; CATALASE_95; 1.
 DR PROSITE: PS00532; CATALASE_96; 1.
 DR PROSITE: PS00533; CATALASE_97; 1.
 DR PROSITE: PS00534; CATALASE_98; 1.
 DR PROSITE: PS00535; CATALASE_99; 1.
 DR PROSITE: PS00536; CATALASE_100; 1.
 DR PROSITE: PS00537; CATALASE_101; 1.
 DR PROSITE: PS00538; CATALASE_102; 1.
 DR PROSITE: PS00539; CATALASE_103; 1.
 DR PROSITE: PS00540; CATALASE_104; 1.
 DR PROSITE: PS00541; CATALASE_105; 1.
 DR PROSITE: PS00542; CATALASE_106; 1.
 DR PROSITE: PS00543; CATALASE_107; 1.
 DR PROSITE: PS00544; CATALASE_108; 1.
 DR PROSITE: PS00545; CATALASE_109; 1.
 DR PROSITE: PS00546; CATALASE_110; 1.
 DR PROSITE: PS00547; CATALASE_111; 1.
 DR PROSITE: PS00548; CATALASE_112; 1.
 DR PROSITE: PS00549; CATALASE_113; 1.
 DR PROSITE: PS00550; CATALASE_114; 1.
 DR PROSITE: PS00551; CATALASE_115; 1.
 DR PROSITE: PS00552; CATALASE_116; 1.
 DR PROSITE: PS00553; CATALASE_117; 1.
 DR PROSITE: PS00554; CATALASE_118; 1.
 DR PROSITE: PS00555; CATALASE_119; 1.
 DR PROSITE: PS00556; CATALASE_120; 1.
 DR PROSITE: PS00557; CATALASE_121; 1.
 DR PROSITE: PS00558; CATALASE_122; 1.
 DR PROSITE: PS00559; CATALASE_123; 1.
 DR PROSITE: PS00560; CATALASE_124; 1.
 DR PROSITE: PS00561; CATALASE_125; 1.
 DR PROSITE: PS00562; CATALASE_126; 1.
 DR PROSITE: PS00563; CATALASE_127; 1.
 DR PROSITE: PS00564; CATALASE_128; 1.
 DR PROSITE: PS00565; CATALASE_129; 1.
 DR PROSITE: PS00566; CATALASE_130; 1.
 DR PROSITE: PS00567; CATALASE_131; 1.
 DR PROSITE: PS00568; CATALASE_132; 1.
 DR PROSITE: PS00569; CATALASE_133; 1.
 DR PROSITE: PS00570; CATALASE_134; 1.
 DR PROSITE: PS00571; CATALASE_135; 1.
 DR PROSITE: PS00572; CATALASE_136; 1.
 DR PROSITE: PS00573; CATALASE_137; 1.
 DR PROSITE: PS00574; CATALASE_138; 1.
 DR PROSITE: PS00575; CATALASE_139; 1.
 DR PROSITE: PS00576; CATALASE_140; 1.
 DR PROSITE: PS00577; CATALASE_141; 1.
 DR PROSITE: PS00578; CATALASE_142; 1.
 DR PROSITE: PS00579; CATALASE_143; 1.
 DR PROSITE: PS00580; CATALASE_144; 1.
 DR PROSITE: PS00581; CATALASE_145; 1.
 DR PROSITE: PS00582; CATALASE_146; 1.
 DR PROSITE: PS00583; CATALASE_147; 1.
 DR PROSITE: PS00584; CATALASE_148; 1.
 DR PROSITE: PS00585; CATALASE_149; 1.
 DR PROSITE: PS00586; CATALASE_150; 1.
 DR PROSITE: PS00587; CATALASE_151; 1.
 DR PROSITE: PS00588; CATALASE_152; 1.
 DR PROSITE: PS00589; CATALASE_153; 1.
 DR PROSITE: PS00590; CATALASE_154; 1.
 DR PROSITE: PS00591; CATALASE_155; 1.
 DR PROSITE: PS00592; CATALASE_156; 1.
 DR PROSITE: PS00593; CATALASE_157; 1.
 DR PROSITE: PS00594; CATALASE_158; 1.
 DR PROSITE: PS00595; CATALASE_159; 1.
 DR PROSITE: PS00596; CATALASE_160; 1.
 DR PROSITE: PS00597; CATALASE_161; 1.
 DR PROSITE: PS00598; CATALASE_162; 1.
 DR PROSITE: PS00599; CATALASE_163; 1.
 DR PROSITE: PS00600; CATALASE_164; 1.
 DR PROSITE: PS00601; CATALASE_165; 1.
 DR PROSITE: PS00602; CATALASE_166; 1.
 DR PROSITE: PS00603; CATALASE_167; 1.
 DR PROSITE: PS00604; CATALASE_168; 1.
 DR PROSITE: PS00605; CATALASE_169; 1.
 DR PROSITE: PS00606; CATALASE_170; 1.
 DR PROSITE: PS00607; CATALASE_171; 1.
 DR PROSITE: PS00608; CATALASE_172; 1.
 DR PROSITE: PS00609; CATALASE_173; 1.
 DR PROSITE: PS00610; CATALASE_174; 1.
 DR PROSITE: PS00611; CATALASE_175; 1.
 DR PROSITE: PS00612; CATALASE_176; 1.
 DR PROSITE: PS00613; CATALASE_177; 1.
 DR PROSITE: PS00614; CATALASE_178; 1.
 DR PROSITE: PS00615; CATALASE_179; 1.
 DR PROSITE: PS00616; CATALASE_180; 1.
 DR PROSITE: PS00617; CATALASE_181; 1.
 DR PROSITE: PS00618; CATALASE_182; 1.
 DR PROSITE: PS00619; CATALASE_183; 1.
 DR PROSITE: PS00620; CATALASE_184; 1.
 DR PROSITE: PS00621; CATALASE_185; 1.
 DR PROSITE: PS00622; CATALASE_186; 1.
 DR PROSITE: PS00623; CATALASE_187; 1.
 DR PROSITE: PS00624; CATALASE_188; 1.
 DR PROSITE: PS00625; CATALASE_189; 1.
 DR PROSITE: PS00626; CATALASE_190; 1.
 DR PROSITE: PS00627; CATALASE_191; 1.
 DR PROSITE: PS00628; CATALASE_192; 1.
 DR PROSITE: PS00629; CATALASE_193; 1.
 DR PROSITE: PS00630; CATALASE_194; 1.
 DR PROSITE: PS00631; CATALASE_195; 1.
 DR PROSITE: PS00632; CATALASE_196; 1.
 DR PROSITE: PS00633; CATALASE_197; 1.
 DR PROSITE: PS00634; CATALASE_198; 1.
 DR PROSITE: PS00635; CATALASE_199; 1.
 DR PROSITE: PS00636; CATALASE_200; 1.
 DR PROSITE: PS00637; CATALASE_201; 1.
 DR PROSITE: PS00638; CATALASE_202; 1.
 DR PROSITE: PS00639; CATALASE_203; 1.
 DR PROSITE: PS00640; CATALASE_204; 1.
 DR PROSITE: PS00641; CATALASE_205; 1.
 DR PROSITE: PS00642; CATALASE_206; 1.
 DR PROSITE: PS00643; CATALASE_207; 1.
 DR PROSITE: PS00644; CATALASE_208; 1.
 DR PROSITE: PS00645; CATALASE_209; 1.
 DR PROSITE: PS00646; CATALASE_210; 1.
 DR PROSITE: PS00647; CATALASE_211; 1.
 DR PROSITE: PS00648; CATALASE_212; 1.
 DR PROSITE: PS00649; CATALASE_213; 1.
 DR PROSITE: PS00650; CATALASE_214; 1.
 DR PROSITE: PS00651; CATALASE_215; 1.
 DR PROSITE: PS00652; CATALASE_216; 1.
 DR PROSITE: PS00653; CATALASE_217; 1.
 DR PROSITE: PS00654; CATALASE_218; 1.
 DR PROSITE: PS00655; CATALASE_219; 1.
 DR PROSITE: PS00656; CATALASE_220; 1.
 DR PROSITE: PS00657; CATALASE_221; 1.
 DR PROSITE: PS00658; CATALASE_222; 1.
 DR PROSITE: PS00659; CATALASE_223; 1.
 DR PROSITE: PS00660; CATALASE_224; 1.
 DR PROSITE: PS00661; CATALASE_225; 1.
 DR PROSITE: PS00662; CATALASE_226; 1.
 DR PROSITE: PS00663; CATALASE_227; 1.
 DR PROSITE: PS00664; CATALASE_228; 1.
 DR PROSITE: PS00665; CATALASE_229; 1.
 DR PROSITE: PS00666; CATALASE_230; 1.
 DR PROSITE: PS00667; CATALASE_231; 1.
 DR PROSITE: PS00668; CATALASE_232; 1.
 DR PROSITE: PS00669; CATALASE_233; 1.
 DR PROSITE: PS00670; CATALASE_234; 1.
 DR PROSITE: PS00671; CATALASE_235; 1.
 DR PROSITE: PS00672; CATALASE_236; 1.
 DR PROSITE: PS00673; CATALASE_237; 1.
 DR PROSITE: PS00674; CATALASE_238; 1.
 DR PROSITE: PS00675; CATALASE_239; 1.
 DR PROSITE: PS00676; CATALASE_240; 1.
 DR PROSITE: PS00677; CATALASE_241; 1.
 DR PROSITE: PS00678; CATALASE_242; 1.
 DR PROSITE: PS00679; CATALASE_243; 1.
 DR PROSITE: PS00680; CATALASE_244; 1.
 DR PROSITE: PS00681; CATALASE_245; 1.
 DR PROSITE: PS00682; CATALASE_246; 1.
 DR PROSITE: PS00683; CATALASE_247; 1.
 DR PROSITE: PS00684; CATALASE_248; 1.
 DR PROSITE: PS00685; CATALASE_249; 1.
 DR PROSITE: PS00686; CATALASE_250; 1.
 DR PROSITE: PS00687; CATALASE_251; 1.
 DR PROSITE: PS00688; CATALASE_252; 1.
 DR PROSITE: PS00689; CATALASE_253; 1.
 DR PROSITE: PS00690; CATALASE_254; 1.
 DR PROSITE: PS00691; CATALASE_255; 1.
 DR PROSITE: PS00692; CATALASE_256; 1.
 DR PROSITE: PS00693; CATALASE_257; 1.
 DR PROSITE: PS00694; CATALASE_258; 1.
 DR PROSITE: PS00695; CATALASE_259; 1.
 DR PROSITE: PS00696; CATALASE_260; 1.
 DR PROSITE: PS00697; CATALASE_261; 1.
 DR PROSITE: PS00698; CATALASE_262; 1.
 DR PROSITE: PS00699; CATALASE_263; 1.
 DR PROSITE: PS00700; CATALASE_264; 1.
 DR PROSITE: PS00701; CATALASE_265; 1.
 DR PROSITE: PS00702; CATALASE_266; 1.
 DR PROSITE: PS00703; CATALASE_267; 1.
 DR PROSITE: PS00704; CATALASE_268; 1.
 DR PROSITE: PS00705; CATALASE_269; 1.
 DR PROSITE: PS00706; CATALASE_270; 1.
 DR PROSITE: PS00707; CATALASE_271; 1.
 DR PROSITE: PS00708; CATALASE_272; 1.
 DR PROSITE: PS00709; CATALASE_273; 1.
 DR PROSITE: PS00710; CATALASE_274; 1.
 DR PROSITE: PS00711; CATALASE_275; 1.
 DR PROSITE: PS00712; CATALASE_276; 1.
 DR PROSITE: PS00713; CATALASE_277; 1.
 DR PROSITE: PS00714; CATALASE_278; 1.
 DR PROSITE: PS00715; CATALASE_279; 1.
 DR PROSITE: PS00716; CATALASE_280; 1.
 DR PROSITE: PS00717; CATALASE_281; 1.
 DR PROSITE: PS00718; CATALASE_282; 1.
 DR PROSITE: PS00719; CATALASE_283; 1.
 DR PROSITE: PS00720; CATALASE_284; 1.
 DR PROSITE: PS00721; CATALASE_285; 1.
 DR PROSITE: PS00722; CATALASE_286; 1.
 DR PROSITE: PS00723; CATALASE_287; 1.
 DR PROSITE: PS00724; CATALASE_288; 1.
 DR PROSITE: PS00725; CATALASE_289; 1.
 DR PROSITE: PS00726; CATALASE_290; 1.
 DR PROSITE: PS00727; CATALASE_291; 1.
 DR PROSITE: PS00728; CATALASE_292; 1.
 DR PROSITE: PS00729; CATALASE_293; 1.
 DR PROSITE: PS00730; CATALASE_294; 1.
 DR PROSITE: PS00731; CATALASE_295; 1.
 DR PROSITE: PS00732; CATALASE_296; 1.
 DR PROSITE: PS00733; CATALASE_297; 1.
 DR PROSITE: PS00734; CATALASE_298; 1.
 DR PROSITE: PS00735; CATALASE_299; 1.
 DR PROSITE: PS00736; CATALASE_300; 1.
 DR PROSITE: PS00737; CATALASE_301; 1.
 DR PROSITE: PS00738; CATALASE_302; 1.
 DR PROSITE: PS00739; CATALASE_303; 1.
 DR PROSITE: PS00740; CATALASE_304; 1.
 DR PROSITE: PS00741; CATALASE_305; 1.
 DR PROSITE: PS00742; CATALASE_306; 1.
 DR PROSITE: PS00743; CATALASE_307; 1.
 DR PROSITE: PS00744; CATALASE_308; 1.
 DR PROSITE: PS00745; CATALASE_309; 1.
 DR PROSITE: PS00746; CATALASE_310; 1.
 DR PROSITE: PS00747; CATALASE_311; 1.
 DR PROSITE: PS00748; CATALASE_312; 1.
 DR PROSITE: PS00749; CATALASE_313; 1.
 DR PROSITE: PS00750; CATALASE_314; 1.
 DR PROSITE: PS00751; CATALASE_315; 1.
 DR PROSITE: PS00752; CATALASE_316; 1.
 DR PROSITE: PS00753; CATALASE_317; 1.
 DR PROSITE: PS00754; CATALASE_318; 1.
 DR PROSITE: PS00755; CATALASE_319; 1.
 DR PROSITE: PS00756; CATALASE_320; 1.
 DR PROSITE: PS00757; CATALASE_321; 1.
 DR PROSITE: PS00758; CATALASE_322; 1.
 DR PROSITE: PS00759; CATALASE_323; 1.
 DR PROSITE: PS00760; CATALASE_324; 1.
 DR PROSITE: PS00761; CATALASE_325; 1.
 DR PROSITE: PS00762; CATALASE_326; 1.
 DR PROSITE: PS00763; CATALASE_327; 1.
 DR PROSITE: PS00764; CATALASE_328; 1.
 DR PROSITE: PS00765; CATALASE_329; 1.
 DR PROSITE: PS00766; CATALASE_330; 1.
 DR PROSITE: PS00767; CATALASE_331; 1.
 DR PROSITE: PS00768; CATALASE_332; 1.
 DR PROSITE: PS00769; CATALASE_333; 1.
 DR PROSITE: PS00770; CATALASE_334; 1.
 DR PROSITE: PS00771; CATALASE_335; 1.
 DR PROSITE: PS00772; CATALASE_336; 1.
 DR PROSITE: PS00773; CATALASE_337; 1.
 DR PROSITE: PS00774; CATALASE_338; 1.
 DR PROSITE: PS00775; CATALASE_339; 1.
 DR PROSITE: PS00776; CATALASE_340; 1.
 DR PROSITE: PS00777; CATALASE_341; 1.
 DR PROSITE: PS00778; CATALASE_342; 1.
 DR PROSITE: PS00779; CATALASE_343; 1.
 DR PROSITE: PS00780; CATALASE_344; 1.
 DR PROSITE: PS00781; CATALASE_345; 1.
 DR PROSITE: PS00782; CATALASE_346; 1.
 DR PROSITE: PS00783; CATALASE_347; 1.
 DR PROSITE: PS00784; CATALASE_348; 1.
 DR PROSITE: PS00785; CATALASE_349; 1.
 DR PROSITE: PS00786; CATALASE_350; 1.
 DR PROSITE: PS00787; CATALASE_351; 1.
 DR PROSITE: PS00788; CATALASE_352; 1.
 DR PROSITE: PS00789; CATALASE_353; 1.
 DR PROSITE: PS00790; CATALASE_354; 1.
 DR PROSITE: PS00791; CATALASE_355; 1.
 DR PROSITE: PS00792; CATALASE_356; 1.
 DR PROSITE: PS00793; CATALASE_357; 1.
 DR PROSITE: PS00794; CATALASE_358; 1.
 DR PROSITE: PS00795; CATALASE_359; 1.
 DR PROSITE: PS00796; CATALASE_360; 1.
 DR PROSITE: PS00797; CATALASE_361; 1.
 DR PROSITE: PS00798; CATALASE_362; 1.
 DR PROSITE: PS00799; CATALASE_363; 1.
 DR PROSITE: PS00800; CATALASE_364; 1.
 DR PROSITE: PS00801; CATALASE_365; 1.
 DR PROSITE: PS00802; CATALASE_366; 1.
 DR PROSITE: PS00803; CATALASE_367; 1.
 DR PROSITE: PS00804; CATALASE_368; 1.
 DR PROSITE: PS00805; CATALASE_369; 1.
 DR PROSITE: PS00806; CATALASE_370; 1.
 DR PROSITE: PS00807; CATALASE_371; 1.
 DR PROSITE: PS00808; CATALASE_372; 1.
 DR PROSITE: PS00809; CATALASE_373; 1.
 DR PROSITE: PS00810; CATALASE_374; 1.
 DR PROSITE: PS00811; CATALASE_375; 1.
 DR PROSITE: PS00812; CATALASE_376; 1.
 DR PROSITE: PS00813; CATALASE_377; 1.
 DR PROSITE: PS00814; CATALASE_378; 1.
 DR PROSITE: PS00815; CATALASE_379; 1.
 DR PROSITE: PS00816; CATALASE_380; 1.
 DR PROSITE: PS00817; CATALASE_381; 1.
 DR PROSITE: PS00818; CATALASE_382; 1.
 DR PROSITE: PS00819; CATALASE_383; 1.
 DR PROSITE: PS00820; CATALASE_384; 1.
 DR PROSITE: PS00821; CATALASE_385; 1.
 DR PROSITE: PS00822; CATALASE_386; 1.
 DR PROSITE: PS00823; CATALASE_387; 1.
 DR PROSITE: PS00824; CATALASE_388; 1.
 DR PROSITE: PS00825; CATALASE_389; 1.
 DR PROSITE: PS00826; CATALASE_390; 1.
 DR PROSITE: PS00827; CATALASE_391; 1.
 DR PROSITE: PS00828; CATALASE_392; 1.
 DR PROSITE: PS00829; CATALASE_393; 1.
 DR PROSITE: PS00830; CATALASE_394; 1.
 DR PROSITE: PS00831; CATALASE_395; 1.
 DR PROSITE: PS00832; CATALASE_396; 1.
 DR PROSITE: PS00833; CATALASE_397; 1.
 DR PROSITE: PS00834; CATALASE_398; 1.
 DR PROSITE: PS00835; CATALASE_399; 1.
 DR PROSITE: PS00836; CATALASE_400; 1.
 DR PROSITE: PS00837; CATALASE_401; 1.
 DR PROSITE: PS00838; CATALASE_402; 1.
 DR PROSITE: PS00839; CATALASE

SQ SEQUENCE 1048 AA; 113701 MW; F5CE8FD309D6883E CRC64;
Query Match 39.5%; Score 45; DB 1; Length 1048;
Best Local Similarity 72.7%; Pred. No. 84;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 LPVPAPVTKDP 11
|||||:|
Db 583 LPVPAPITVPP 593

Search completed: May 20, 2003, 16:35:39
Job time : 7.65385 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 25.9615 Seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-15
Perfect score: 114
Sequence: 1 LPVPAPVTKDPATSLYTIPIFH 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	70.2	419	10	08SI10
2	66	57.9	424	10	08SI10
3	53	46.5	173	16	08XU05
4	52	45.6	422	10	08SI03
5	52	45.6	492	12	08OVD2
6	52	45.6	991	12	09I135
7	52	45.6	1435	12	09DS00
8	51.5	45.2	100	12	09DS00
9	51	44.7	442	10	08SI06
10	51	44.7	555	2	08SI737
11	50	43.9	349	11	0923B9
12	50	43.9	434	10	08SI05
13	50	43.9	434	10	08SI04
14	50	43.9	496	4	096SL8
15	50	43.9	2037	13	09PV22
16	50	43.9	3498	5	020497

17	49.5	43.4	726	12	08V7E7	08V7E7 tt virus. O
18	49	43.0	96	16	08XNB8	08XNB8 ralsionia s
19	49	43.0	532	16	P72605	P72605 synechocyst
20	49	43.0	725	4	09H986	09H986 homo sapien
21	49	43.0	835	4	09B011	09B011 homo sapien
22	49	43.0	1011	16	0981K8	0981K8 rhizobium 1
23	49	43.0	1047	5	024019	024019 drosophila
24	49	43.0	1459	10	0941A8	0941A8 arabidopsis
25	48.5	42.5	402	10	09SY96	09SY96 arabidopsis
26	48	42.1	147	10	049128	049128 fitillaria
27	48	42.1	147	10	049129	049129 fitillaria
28	48	42.1	147	10	049130	049130 fitillaria
29	48	42.1	147	10	049131	049131 fitillaria
30	48	42.1	147	10	050053	050053 fitillaria
31	48	42.1	147	10	09S7X1	09S7X1 fitillaria
32	48	42.1	273	16	09KNT2	09KNT2 vibrio chol
33	48	42.1	383	13	091917	091917 brachydanio
34	48	42.1	434	10	09ZV85	09ZV85 arabidopsis
35	48	42.1	753	16	09RD19	09RD19 streptomyce
36	47	41.2	107	10	09FXM1	09FXM1 lilium long
37	47	41.2	336	5	018839	018839 caenorhabdi
38	47	41.2	336	16	055451	055451 synechocyst
39	47	41.2	344	11	089037	089037 fatus norv
40	47	41.2	687	5	0967R6	0967R6 aplysia cal
41	47	41.2	750	16	08Z6H1	08Z6H1 salmonella
42	47	41.2	750	16	09L328	09L328 salmonella
43	47	41.2	963	5	09VYS2	09VYS2 drosophila
44	47	41.2	1199	11	054774	054774 mus musculu
45	46	40.4	186	16	006295	006295 mycobacteri

ALIGNMENTS

RESULT 1	08SI10	PRELIMINARY;	PRT;	419 AA.
ID	08SI10			
AC	08SI10:			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Putative dermal glycoprotein.			
GN	P0504E02.9.			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. NIPPONBARE;			
RA	Sasaki T., Matsunoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC			
RT	clone:P0504E02."			
RL	Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AP003269; BAB89708.1; -			
SQ	SEQUENCE 419 AA: 43786 MW; 792B5F5A4F33A8 CRC64;			
Query Match	70.2%;	Score 80;	DB 10;	Length 419;
Best Local Similarity	83.3%;	Pred. No. 0.0014;		
Matches 15;	Conservative	1;	Mismatches 2;	Indels 0;
Gaps				
OY	2 PVPAPVTKDPATSLYTIIP 19			
DB	37 PVPAPVTKDPATSLYTIIP 54			
RESULT 2	08SI10	PRELIMINARY;	PRT;	424 AA.
ID	08SI10			
AC	08SI10:			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative dermal glycoprotein.
 CN P0504E02.8.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0504E02."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003269; BAB9707.1;
 SO SEQUENCE 424 AA; 44620 MW; 0DACD5CAC852C34 CRC64;

Query Match 57.9%; Score 66; DB 10; Length 424;
 Best Local Similarity 66.7%; Pred. No. 0.15;
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 PVPAPVTKDPATSLYTI 19
 DB 34 PLVAATTKDPAATSLYTPV 51

RESULT 3
 ID 08XUH5 PRELIMINARY; PRT; 173 AA.
 AC 08XUH5;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Hypothetical protein RSC3213.
 GN RSC3213 OR RSC02430.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangot S.,
 Arian M., Billault A., Brotier P., Camus J.C., Catolico L.,
 Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 Sigler P., Thebault P., Whalen M., Wincker P., Levy W.,
 Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 KW EMBL: AL646074; CADI7001.1;
 DR Hypothetical protein; Complete proteome.
 SO SEQUENCE 173 AA; 18665 MW; F902A56F9D25E737 CRC64;

Query Match 46.5%; Score 53; DB 16; Length 173;
 Best Local Similarity 64.3%; Pred. No. 4.7;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSL 15
 DB 72 PAPAPVTEDEPAAEV 85
 RESULT 4
 ID 08S1V3 PRELIMINARY; PRT; 422 AA.
 AC 08S1V3;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative dermal glycoprotein.

GN P0504E02.6.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0504E02."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003269; BAB9705.1;
 SO SEQUENCE 422 AA; 44745 MW; 08CFE26346400A9F CRC64;

Query Match 45.6%; Score 52; DB 10; Length 422;
 Best Local Similarity 64.7%; Pred. No. 16;
 Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 2 PVPAPVTKDPATSLYTI 18
 DB 37 PLVTAATKKGATKLYTI 53

RESULT 5
 ID 08QVD2 PRELIMINARY; PRT; 492 AA.
 AC 08QVD2;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE RNA-dependent RNA polymerase (Fragment).
 OS Colorado tick fever virus.
 OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.
 OX NCBI_TaxID=46839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CTFV-S6-14-03;
 RX MEDLINE=21954629; PubMed=11958454;
 RA Attoui H., Mohd J.F., Biagini P., Cantaloube J.F., de M.P.,
 Murphy F.A., de L.X.;
 RT "Genus Coltivirus (family Reoviridae): genomic and morphologic
 characterization of Old World viruses";
 RL Arch. Virol. 147:533-561(2002).
 DR EMBL: AF343051; AAM18357.1;
 KW RNA-directed RNA polymerase.
 FT NON_TER 1 492 1
 FT NON_TER 1 492 1
 SO SEQUENCE 492 AA; 55966 MW; 501250E88F84DA58 CRC64;

Query Match 45.6%; Score 52; DB 12; Length 492;
 Best Local Similarity 52.2%; Pred. No. 19;
 Matches 12; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 LPVPAPVTKD-----PATSLYTI 19
 DB 408 VPPPLVOKDELICPGTSVYTI 430
 RESULT 6
 ID 099135 PRELIMINARY; PRT; 991 AA.
 AC 099135;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE RNA polymerase (Fragment).
 OS Colorado tick fever virus.
 OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.
 OX NCBI_TaxID=46839;
 RN [1]
 RP SEQUENCE FROM N.A.

ID 096SL8 PRELIMINARY; PRT: 496 AA.
 AC 096SL8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CDNA FLJ14768 fis, clone NT2RP3004125, moderately similar to Mus musculus zinc finger protein splice variant FIZ1-B (FIZ1) mRNA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Ninomiya K., Iwayanagi T.,
 RA "NEDO human cDNA sequencing project."
 RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK027674; BAB55286.1;
 DR InterPro: IPR000564; 2FE2S_Ferredoxin.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF00096; Zf-C2H2; 1.
 DR ProDom: PD000003; Znf_C2H2; 1.
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_10.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 DR DNA-binding: Zinc-finger.
 SQ SEQUENCE 496 AA; 51993 MW; 937EEC713FFBD58F CRC64;

Query Match 43.9%; Score 50; DB 4; Length 496;
 Best Local Similarity 38.1%; Pred. No. 37;
 Matches 8; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 LPVPAPVTRKPAATSLYTIPIF 21
 :| | | | | : |||
 DB 4 VPAPTRPAPAPAAAPRVPIF 24

RESULT 15
 09PVZ2 PRELIMINARY; PRT: 2037 AA.
 ID 09PVZ2;
 AC 09PVZ2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Nucleoporin CAN.
 GN Nucleoporin CAN.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-99384295; PubMed-10454574;
 RA Askjaer P., Bachl A., Wilm M., Bischoff R., Weeks D.L., Ogniewski V., Ohno M., Niehrs C., Kjems J., Matla J.W., Fornerod M.,
 RT "RangRP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase."
 RL Mol. Cell. Biol. 19:6276-6285(1999).
 DR EMBL: AJ243889; CAB53357.1; -;
 DR InterPro: IPR000515; BPD_transp.
 DR InterPro: IPR004325; Nucleoporin_FG.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF03093; Nucleoporin_FG; 16.
 DR Pfam: PF00400; WD40; 2.
 DR SMART; SM00320; WD40; 2.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; UNKNOWN_1.
 KW Porin; Repeat; WD repeat.

SQ SEQUENCE 2037 AA; 208932 MW; 2250EA49A0435635 CRC64;
 Query Match 43.9%; Score 50; DB 13; Length 2037;
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 5 APYTKDPATSLYTIPIF 20
 ||: | |||| | ||
 DB 1672 APLMKGPATSGRIPIF 1687

Search completed: May 20, 2003, 16:43:13
 Job time : 28.9615 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 9.69231 Seconds
(without alignments)
63,750 Million cell updates/sec

Title: US-09-869-155-15

Sequence: 1 LPVPAPYTKDPATSLYRPFH 21

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	39.5	690	US-08-971-089-6	Sequence 6, Appli
2	45	39.5	822	US-09-564-805-222	Sequence 222, App
3	44	38.6	353	US-08-073-807A-17	Sequence 17, Appl
4	44	38.6	380	US-08-073-807A-18	Sequence 18, Appl
5	44	38.6	416	US-08-073-807A-2	Sequence 2, Appli
6	43.5	38.2	978	US-08-415-593-43	Sequence 43, Appl
7	43	37.7	288	US-09-651-200-14	Sequence 14, Appl
8	43	37.7	292	US-09-651-200-16	Sequence 16, Appl
9	43	37.7	297	US-08-580-545B-6	Sequence 6, Appli
10	43	37.7	322	US-09-262-653A-6	Sequence 6, Appli
11	43	37.7	322	US-08-892-880-2	Sequence 2, Appli
12	43	37.7	322	US-09-232-160-21	Sequence 21, Appl
13	43	37.7	751	US-09-036-987A-24	Sequence 24, Appl
14	43	37.7	751	US-09-370-700-24	Sequence 24, Appl
15	42.5	37.3	432	US-08-749-902-6	Sequence 6, Appli
16	42	36.8	135	US-09-134-001C-3069	Sequence 3069, Ap
17	42	36.8	383	US-09-134-001C-3068	Sequence 3068, Ap
18	42	36.8	693	US-08-553-279-2	Sequence 2, Appli
19	42	36.8	826	US-09-564-805-2	Sequence 2, Appli
20	42	36.8	826	US-09-564-805-224	Sequence 224, Appl
21	42	36.8	846	US-09-564-805-226	Sequence 226, App
22	42	36.8	1018	US-09-134-001C-3067	Sequence 3067, Ap
23	42	36.8	1018	US-08-072-610-2	Sequence 2, Appli
24	42	36.8	1018	US-08-719-822B-2	Sequence 2, Appli
25	42	36.8	1018	US-09-092-458-2	Sequence 2, Appli
26	42	36.8	1183	US-09-134-001C-3530	Sequence 3530, Ap
27	42	36.8	2182	US-08-487-826B-16	Sequence 16, Appl

28	41.5	36.4	29	US-08-945-168-103	Sequence 103, App
29	41.5	36.4	379	US-09-320-878-6	Sequence 6, Appli
30	41.5	36.4	379	US-09-105-537-16	Sequence 16, Appli
31	41.5	36.4	488	US-07-794-393-2	Sequence 2, Appli
32	41.5	36.4	488	US-08-001-711-2	Sequence 2, Appli
33	41.5	36.4	488	US-08-704-711A-22	Sequence 22, Appli
34	41.5	36.4	488	US-09-521-220-22	Sequence 22, Appli
35	41.5	36.4	488	US-09-391-104-31	Sequence 31, Appli
36	41.5	36.4	489	US-08-448-489-11	Sequence 11, Appli
37	41.5	36.4	3782	US-09-105-537-4	Sequence 4, Appli
38	41	36.0	112	US-09-134-001C-5057	Sequence 5057, Ap
39	41	36.0	178	US-08-928-443-1	Sequence 1, Appli
40	41	36.0	178	US-09-129-055-1	Sequence 1, Appli
41	41	36.0	279	US-08-872-979-5	Sequence 5, Appli
42	41	36.0	289	US-08-393-985-8	Sequence 8, Appli
43	41	36.0	289	US-08-872-979-6	Sequence 6, Appli
44	41	36.0	310	US-08-796-792-4	Sequence 4, Appli
45	41	36.0	480	US-09-108-020-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-971-089-6
; Sequence 6, Application US/08971089
; Patent No. 6376174
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; APPLICANT: Scoles, Daniel R.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING
; TITLE OF INVENTION: SCHWANNOMIN-BINDING-PROTEINS AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.089
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,987
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramos, Robert T.
; REGISTRATION NUMBER: 37,915
; REFERENCE/DOCKET NUMBER: P-CE 2862
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-971-089-6
Query Match 39.5%; Score 45; DB 4; Length 690;
Best Local Similarity 44.4%; Pred. No. 69;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 PVPAPYTKDPATSLYRPF 19
|:|:|:|:|:|:|

Db 294 PMPSSASAPPASSLYSP 311

RESULT 2

US-09-564-805-222
Sequence 222, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/564,805
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRICR APPLICATION NUMBER: 09/434,382
PRICR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 222
LENGTH: 822
TYPE: PRT
ORGANISM: Mus musculus
US-09-564-805-222

Query Match: 39.5%; Score 45; DB 4; Length 822;
Best Local Similarity 40.0%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTI 21

Db 159 PHSAPRYKDEMTYQVPIH 178

RESULT 3

US-08-073-807A-17
Sequence 17, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
APPLICANT: Fukuda, Minoru
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,807A
FILING DATE: 08-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9567
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-073-807A-17

Query Match: 38.6%; Score 44; DB 1; Length 353;
Best Local Similarity 41.2%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTI 18

Db 180 PPSPPVPKSPSPVDKYNV 196

RESULT 4

US-08-073-807A-18
Sequence 18, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
APPLICANT: Fukuda, Minoru
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,807A
FILING DATE: 08-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9567
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-073-807A-18

Query Match: 38.6%; Score 44; DB 1; Length 380;
Best Local Similarity 41.2%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTI 18

Db 207 PPSPPVPKSPSPVDKYNV 223

RESULT 5

US-08-073-807A-2
Sequence 2, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.

```

1  APPLICANT : Fukuda, CELU
2  TITLE OF INVENTION : CELL SURFACE LAMP EXPRESSION AND
3  TITLE OF INVENTION : SELECTIN-DEPENDENT ADHESION
4  NUMBER OF SEQUENCES : 18
5  CORRESPONDENCE ADDRESS :
6  ADDRESSEE : Campbell and Flores
7  STREET : 4370 La Jolla Village Drive, Suite 700
8  CITY : San Diego
9  STATE : California
10 COUNTRY : United States
11 ZIP : 92122
12
13 COMPUTER READABLE FORM :
14 MEDIUM TYPE : Floppy disk
15 COMPUTER : IBM PC compatible
16 OPERATING SYSTEM : PC-DOS/MS-DOS
17 SOFTWARE : Patent In Release #1.0, Version #1.25
18
19 CURRENT APPLICATION DATA :
20 APPLICATION NUMBER : US/08/073,807A
21 FILING DATE : 08-JUN-1993
22 CLASSIFICATION : -435
23
24 ATTORNEY/AGENT INFORMATION :
25 NAME : Campbell, Cathryn A.
26 REGISTRATION NUMBER : 31,815
27 REFERENCE/DOCKET NUMBER : P-LJ 9567
28 TELECOMMUNICATION INFORMATION :
29 TELEPHONE : (619) 535-9001
30 TELEFAX : (619) 535-8949
31
32 INFORMATION FOR SEQ ID NO : 2 :
33 SEQUENCE CHARACTERISTICS :
34 LENGTH : 416 amino acids
35 TYPE : amino acid
36 TOPOLOGY : linear
37
38 MOLECULE TYPE : protein
39
40 US-08-073-807A-2

```

Query Match	38.6%	Score 44:	DB 1,	Length 416,
Best Local Similarity	41.2%	Pred. No. 53,		
Matches 7, Conservative	3;	Mismatches 7;	Indels 0;	Gaps 0;

```
QY      2 PVPAPVTKDPATSLYTI 18
          | | : | | | : | :
Db      207 PSPSPVPKSPSPVDKYNV 223
```

US-08-415-593-43
Sequence 43, Application US/08415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Nianjun
APPLICANT: Politis-Virk, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millita Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:

```

? CLASSIFICATION: 4355
? ATTORNEY/AGENT INFORMATION:
? NAME: BROOK, David E.
? REGISTRATION NUMBER: 22,592
? REFERENCE/DOCKET NUMBER: CPI94-09
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-861-6240
? TELEFAX: 617-862-6540
? INFORMATION FOR SEQ ID NO: 43:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 978 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-08-415-593-43

```

Query Match	38.2%	Score 43.5;	DB 2;	Length 978;
Best Local Similarity	39.1%;	Pred. No. 1.8e+02;		
Matches	9;	Conservative	5;	Mismatches 2;
			Indels	7;
			Gaps	1;

```
QY      6  PVTKDPATSL-----YTIPFH 21
          | : : | : | | | : | : | |
Db     116 PLEDPSTSLAWTTTPWTLP SH 138
```

RESULT 7
US-09-651-200-14
; Sequence 14, Application US/09651200
; Patent No. 6429303

? APPLICANT: Green et al
 ? TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 ? TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 ? TITLE OF INVENTION: Polypeptides Encoded Thereby
 ? FILE REFERENCE: 15966-562 (CURA-62)
 ? CURRENT APPLICATION NUMBER: 05/09/651,200
 ? CURRENT FILING DATE: 2000-08-30
 ? PRIOR APPLICATION NUMBER: 60/152383
 ? PRIOR FILING DATE: 1999-09-03
 ? PRIOR APPLICATION NUMBER: 60/172909
 ? PRIOR FILING DATE: 1999-12-21
 ? PRIOR APPLICATION NUMBER: 60/183578
 ? PRIOR FILING DATE: 2000-02-18
 ? NUMBER OF SEQ ID NOS: 25
 ? SOFTWARE: PatentIn Ver. 2.0
 ? SEQ ID NO 14
 ? LENGTH: 288
 ? TYPE: PRT
 ? ORGANISM: Macaca mulatta
 ? US-09-651-200-14

Query Match	37.7%;	Score 43;	DB 4;	Length 288;
Best Local Similarity	43.8%;	Pred. No. 48;		
Matches	7;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0;

Qy 3 VPAPVTKDPATSLYTI 18
 : |::|||
 Db 185 ISTVVSQDPETELTYTV 200

```

; RESULT 8
; US-09-651-200-16
; Sequence 16, Application US/09651200
; Patent No. 6429303
; GENERAL INFORMATION:
; APPLICANT: Green et al
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-562 (CURA-62)
; CURRENT APPLICATION NUMBER: US/09/651,200
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152363

```



```

APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merio, Donald J.
APPLICANT: Treadway, Patli J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-24

Query Match      37.7% Score 43; DB 4; Length 751;
Best Local Similarity 31.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0.

QY          3 VPAPVTKDPATSLYTIIPFH 21
           :|::|||
Db          282 IPTDDDEVPTVAIIVPFH 300

RESULT 14
US-09-370-700-24
Sequence 24, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patli J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 24
LENGTH: 751
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
```

US-09-370-700-24

Query Match 37.7%; Score 43; DB 4; Length 751;

Best Local Similarity 31.6%; Pred. No. 1.6e+02;

Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 VPAPVTKDPATSLVTYPE 21

DB 282 IPTDDDEVPTVAIYLVPFH 300

RESULT 15

US-08-749-902-6

Sequence 6, Application US/08749902

Patent No. 5985635

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Goll, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE

TITLE OF INVENTION: PROTEIN KINASES

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEO Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/749,902

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0150 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1016651

US-08-749-902-6

Query Match 37.3%; Score 42.5; DB 2; Length 432;

Best Local Similarity 42.9%; Pred. No. 94;

Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 LPV-PAPVTKDPATSLVTYPE 20

DB 323 VPPIPSPTKDRKRSGLTVPY 343

Search completed: May 20, 2003, 16:44:40
Job time: 11.6923 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 ; Search time 10.3846 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-15
Perfect score: 114
Sequence: 1 LPVPAPYTKDPATSLYIPFH 21

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	42.1	81	10 US-09-864-761-47685	Sequence 47685, A
2	47	41.2	750	10 US-09-815-242-13985	Sequence 13985, A
3	46	40.4	2701	9 US-10-171-311-83	Sequence 83, Appl
4	45	39.5	136	9 US-10-091-483-211	Sequence 211, App
5	45	39.5	136	10 US-09-764-846-211	Sequence 211, App
6	45	39.5	635	9 US-09-738-626-6614	Sequence 6614, App
7	45	39.5	690	9 US-10-117-604-6	Sequence 6, Appli
8	45	39.5	753	10 US-09-815-242-10181	Sequence 10181, A
9	45	39.5	822	9 US-09-988-626-222	Sequence 222, App
10	45	39.5	822	9 US-09-988-667-222	Sequence 222, App
11	45	39.5	1708	9 US-09-468-147-166	Sequence 166, App
12	44.5	39.0	532	9 US-10-225-928-2	Sequence 2, Appli
13	44	38.6	475	10 US-09-927-091-1	Sequence 21, Appl
14	44	38.6	504	8 US-08-980-068B-21	Sequence 21, Appl
15	43.5	38.2	328	9 US-10-243-739-11	Sequence 11, Appl
16	43.5	38.2	375	12 US-10-062-254-373	Sequence 373, App
17	43.5	38.2	580	10 US-09-808-387-36	Sequence 36, Appl
18	43	37.7	138	10 US-09-925-301-1164	Sequence 1164, Ap
19	43	37.7	156	9 US-09-907-969-582	Sequence 582, App

20	43	37.7	292	10 US-09-303-510-2	Sequence 2, Appli
21	43	37.7	292	10 US-09-303-510-4	Sequence 4, Appli
22	43	37.7	292	10 US-09-303-040-2	Sequence 2, Appli
23	43	37.7	292	10 US-09-303-040-4	Sequence 4, Appli
24	43	37.7	294	9 US-09-907-969-487	Sequence 487, App
25	43	37.7	294	10 US-09-884-441-487	Sequence 487, App
26	43	37.7	322	9 US-09-905-291A-201	Sequence 201, App
27	43	37.7	322	9 US-10-063-547-6	Sequence 6, Appli
28	43	37.7	322	9 US-09-902-853-201	Sequence 201, App
29	43	37.7	322	9 US-09-907-824-201	Sequence 201, App
30	43	37.7	322	9 US-09-907-841-201	Sequence 201, App
31	43	37.7	322	9 US-09-904-011-201	Sequence 201, App
32	43	37.7	322	9 US-10-063-616-6	Sequence 6, Appli
33	43	37.7	322	9 US-09-906-742-201	Sequence 201, App
34	43	37.7	322	9 US-10-063-502-6	Sequence 6, Appli
35	43	37.7	322	9 US-09-908-838-201	Sequence 201, App
36	43	37.7	322	9 US-09-907-613-201	Sequence 201, App
37	43	37.7	322	9 US-09-907-942-201	Sequence 201, App
38	43	37.7	322	9 US-09-904-820-201	Sequence 201, App
39	43	37.7	322	9 US-09-904-859-201	Sequence 201, App
40	43	37.7	322	9 US-09-909-204-201	Sequence 201, App
41	43	37.7	322	9 US-09-904-786-201	Sequence 201, App
42	43	37.7	322	9 US-09-906-646-201	Sequence 201, App
43	43	37.7	322	9 US-09-906-700-201	Sequence 201, App
44	43	37.7	322	9 US-09-902-903-201	Sequence 201, App
45	43	37.7	322	9 US-09-903-749A-201	Sequence 201, App

ALIGNMENTS

RESULT 1

US-09-864-761-47685
Sequence 47685, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 47685
: LENGTH: 81
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC009305.1
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.57
: OTHER INFORMATION: EST_HUMAN HIT: BE881367.1, EVALUO 3.00e-04
: OTHER INFORMATION: SWISSPROT HIT: Q09472, EVALUO 2.00e-04
US-09-364-761-47685
```

```

Query Match          42.1%; Score 48; DB 10; Length 81;
Best Local Similarity 45.0%; Pred. No. 7.8;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

```
QY  .1 LPVPAPVTKDPATSLVYTF 20
    ||: ||: ||: ||: ||: ||
Db  48 LPATPVSQPPQTSIDSVTF 67
```

```

RESULT 2
US-09-815-242-13985
: Sequence 13985, Application US/09815242
: Patent No. US20020061569A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes In
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13985
: LENGTH: 750
: TYPE: PRT
: ORGANISM: Salmonella typhi
US-09-815-242-13985
```

```

Query Match          41.2%; Score 47; DB 10; Length 750;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
```

```

QY  4 PAP---VTKDPATSLVYTF 19
    ||: ||: ||: ||: ||: ||
Db  572 PPPDVNGLKDPALSLVAVP 591
```

```

RESULT 3
US-10-171-311-83
: Sequence 83, Application US/10171311
: Publication No. US20030087270A1
: GENERAL INFORMATION:
```

```

: APPLICANT: Schlegel, Robert
: APPLICANT: Chen, Yan
: APPLICANT: Zhao, Xumei
: APPLICANT: Monahan, John
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Gialt, Karen
: APPLICANT: Gannavarapu, Manjula
: APPLICANT: Hoerish, Sebastian
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
: FILE REFERENCE: MRI-035
: CURRENT APPLICATION NUMBER: US/10/171,311
: PRIOR FILING DATE: 2002-06-12
: PRIOR APPLICATION NUMBER: US 60/298,159
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,155
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/335,936
: PRIOR FILING DATE: 2001-11-14
: NUMBER OF SEQ ID NOS: 238
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 83
: LENGTH: 2701
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-171-311-83
```

```

Query Match          40.4%; Score 46; DB 9; Length 2701;
Best Local Similarity 44.4%; Pred. No. 8.1e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
```

```

QY  2 PVPAPVTKDPATSLVYTF 19
    ||| ||: ||: ||: ||: ||
Db  1874 PTPAPASSPAPVITAP 1891
```

```

RESULT 4
US-10-091-483-211
: Sequence 211, Application US/10091483
: Publication No. US20030049650A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT212C1
: CURRENT APPLICATION NUMBER: US/10/091,483
: CURRENT FILING DATE: 2002-03-07
: NUMBER OF SEQ ID NOS: 348
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 211
: LENGTH: 136
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-091-483-211
```

```

Query Match          39.5%; Score 45; DB 9; Length 136;
Best Local Similarity 57.1%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```

QY  1 LPVPAPVTKDPATS 14
    ||: ||: ||: ||: ||: ||
```


Db 76 VPLPVPDTKPPPTS 89

RESULT 5
US-09-764-846-211

; Sequence 211, Application US/09764846
; Patent No. US20020102638A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT212

; CURRENT APPLICATION NUMBER: US/09/764,846

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 348

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 211

; LENGTH: 136

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-846-211

Query Match 39.5%; Score 45; DB 10; Length 136;
Best Local Similarity 57.1%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LPVPAPVTKDPATPS 14

Db 76 VPLPVPDTKPPPTS 89

RESULT 6

US-09-738-626-6614

; Sequence 6614, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAMA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAOKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO: 6614

; LENGTH: 635

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-6614

Query Match 39.5%; Score 45; DB 9; Length 635;
Best Local Similarity 50.0%; Pred. No. 2,1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATPSLYTIP 19

Db 181 PVPAPVTSAPGASVPSIP 198

RESULT 7

US-10-117-604-6

; Sequence 6, Application US/10117604

; Patent No. US20020168672A1

; GENERAL INFORMATION:

; APPLICANT: Pulst, Stefan M.

; APPLICANT: Sciles, Daniel R.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING

; SCHWANNOMIN-BINDING-PROTEINS AND PRODUCTS RELATED THEREIN

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESSES:

; ADDRESS: CAMPBELL & FLORES, LLP

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/117,604

; FILING DATE: 04-Apr-2002

; CLASSIFICATION: <unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/971,089

; FILING DATE: <unknown>

; APPLICATION NUMBER: US 60/030,987

; FILING DATE: 15-Nov-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Ramos, Robert T.

; REGISTRATION NUMBER: 37,915

; REFERENCE/DOCKET NUMBER: P-CE 2862

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619)535-9001

; TELEFAX: (619)535-8949

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 690 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-117-604-6

Query Match 39.5%; Score 45; DB 9; Length 690;
Best Local Similarity 44.4%; Pred. No. 2.4e+02;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATPSLYTIP 19

Db 294 PMSASSAPPASSLYSSP 311

RESULT 8

US-09-815-242-10181

; Sequence 10181, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10181
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Escherichia coli
US-05-815-242-10181

```

```

Query Match          39.5%; Score 45; DB 10; Length 753;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

```

```
QY 4 PAP---VTKDPATSLYTIPT 19
      |||  |||  |||  |||  |||
DB 575 PPPDVNGLKDPDSLAIPT 594

```

RESULT 9

```

; Sequence 222, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavligian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-988-626-222

```

```

Query Match          39.5%; Score 45; DB 9; Length 822;
Best Local Similarity 40.0%; Pred. No. 2.9e+02;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```

```
QY 2 PVPAPVTKDPATSLYTIPT 21
      |||  |||  |||  |||  |||
DB 159 PHSAPPEYKDEMTVYQVPIH 178

```

RESULT 10

```

US-09-988-687-222
; Sequence 222, Application US/09988687

```

```

; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavligian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-988-687-222

```

```

Query Match          39.5%; Score 45; DB 9; Length 822;
Best Local Similarity 40.0%; Pred. No. 2.9e+02;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```

```
QY 2 PVPAPVTKDPATSLYTIPT 21
      |||  |||  |||  |||  |||
DB 159 PHSAPPEYKDEMTVYQVPIH 178

```

RESULT 11

```

; Sequence 166, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Etker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; FILE REFERENCE: 6232-US-P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 1708
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 322
; OTHER INFORMATION: Xaa = Unknown or Other at position 331
; OTHER INFORMATION: Xaa = Unknown or Other at position 445
; OTHER INFORMATION: Xaa = Unknown or Other at position 448
; OTHER INFORMATION: Xaa = Unknown or Other at position 634
; OTHER INFORMATION: Xaa = Unknown or Other at position 646
; OTHER INFORMATION: Xaa = Unknown or Other at position 811
; OTHER INFORMATION: Xaa = Unknown or Other at position 1533
; OTHER INFORMATION: Xaa = Unknown or Other at position 1578
; OTHER INFORMATION: Xaa = Unknown or Other at position 1691
US-09-468-147-166

```

Query Match 39.5%; Score 45; DB 9; Length 1708;
Best Local Similarity 42.3%; Pred. No. 6.6e+02;
Matches 11; Conservative 1; Mismatches 6; Indels 8; Gaps 1;

QY 2 PVPAPVTKDPATSLYTIPTF 19
DB 775 PPPPPVPRKPPATSPPPRTTLLTYTP 800

RESULT 12
US-10-225-928-2

; Sequence 2, Application US/10225928
; Publication No. US20030092041A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; TITLE OF INVENTION: NOVEL USE FOR MUSCARINIC RECEPTOR M5 IN
; FILE REFERENCE: MPI01-210PIRM
; CURRENT APPLICATION NUMBER: US/10/225,928
; CURRENT FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: human
US-10-225-928-2

Query Match 39.0%; Score 44.5; DB 9; Length 532;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 1 LPVPAPVTKDPATSLYTIPTF 21
DB 403 MCPPEPAKEPSTKGLNPNPSH 424

RESULT 13

US-09-927-091-1
; Sequence 1, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOFT, STEVE
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:6510S
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Human
US-09-927-091-1

Query Match 38.6%; Score 44; DB 10; Length 475;
Best Local Similarity 61.5%; Pred. No. 2.1e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVPAPVTKDPATSLYTIPTF 14
DB 292 PVPALTLDPCTA 304

RESULT 14
US-08-980-068B-21

; Sequence 21, Application US/08980068B
; Patent No. US20020081718A1
; GENERAL INFORMATION:
; APPLICANT: HOTTA, Yoshiaki
; TITLE OF INVENTION: A POLYPEPTIDE COMMON TO GLIAL CELLS MISSING (GCM)
; FILE REFERENCE: 97-1513*/LC/00653
; CURRENT APPLICATION NUMBER: US/08/980,068B
; CURRENT FILING DATE: 1997-11-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Drosophila
US-08-980-068B-21

Query Match 38.6%; Score 44; DB 8; Length 504;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LPVPAPVTKDPATSLYTIPTF 13
DB 14 MPVPMVPSPAT 26

RESULT 15
US-10-243-739-11

; Sequence 11, Application US/10243739
; Publication No. US20030091593A1
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Bacteriophage Q-beta
US-10-243-739-11

Query Match 38.2%; Score 43.5; DB 9; Length 328;
Best Local Similarity 45.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 LPVPAPVTKDPATSLYTIPTF 20
DB 150 IPDP-PIDPEPCKWYTCPF 168

Search completed: May 20, 2003, 18:00:01
Job time: 11.3846 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 ; Search time 33.7253 seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-16

Perfect score: 145

Sequence: 1 ILASLPKSTGVAGLANSGLAPQAVASAK 31

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	31	21	AAV93759
2	145	100.0	31	22	AAE14690
3	145	100.0	31	22	AAU07395
4	145	100.0	370	23	AAU75815
5	145	100.0	370	23	AAU75816
6	145	100.0	381	23	AAU75813
7	140	96.6	381	23	AAU75812
8	138	95.2	381	23	AAU75814
9	125	86.2	277	23	AAU75818
10	111	76.6	287	23	AAU75819

11	78	53.8	185	23	AAU75824
12	78	53.8	185	23	AAU75825
13	78	53.8	185	23	AAU75826
14	64	44.1	433	21	AAU75829
15	64	44.1	433	21	AAU75830
16	64	44.1	433	21	AAU75831
17	64	44.1	433	21	AAU75832
18	64	44.1	433	21	AAU75833
19	59	40.7	249	21	AAU75834
20	57	39.3	386	21	AAU75835
21	55	37.9	141	23	AAU75836
22	55	37.9	95	23	AAU75837
23	55	37.9	338	22	AAU75838
24	55	37.9	381	23	AAU75839
25	55	37.9	394	8	AAU75840
26	52	35.9	340	11	AAU75841
27	50.5	34.8	374	21	AAU75842
28	50	34.5	1114	23	AAU75843
29	49.5	34.1	217	22	AAU75844
30	49.5	34.1	716	22	AAU75845
31	49.5	34.1	722	22	AAU75846
32	49	33.8	104	22	AAU75847
33	49	33.8	290	22	AAU75848
34	49	33.8	322	23	AAU75849
35	49	33.8	326	23	AAU75850
36	49	33.8	326	23	AAU75851
37	49	33.8	1010	22	AAU75852
38	49	33.8	2237	22	AAU75853
39	48.5	33.4	71	19	AAU75854
40	48.5	33.4	1225	22	AAU75855
41	48.5	33.4	1258	22	AAU75856
42	48.5	33.4	1429	22	AAU75857
43	48	33.1	152	22	AAU75858
44	48	33.1	208	14	AAU75859
45	48	33.1	214	22	AAU75860

ALIGNMENTS

RESULT 1
ID AAV93759 standard; peptide: 31 AA.
AC AAV93759:
DT 03-OCT-2000 (first entry)
DE Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.
KW Endo-beta-1,4-xylanase inhibitor: wheat flour; xylanase; bakery product;
KW dough: dough preparation.
OS Triticum sp.
PN WO200039289-A2.
PD 06-JUL-2000.
PF 17-DEC-1999; 99WO-IB02071.
PR 23-DEC-1998; 98GB-0028599.
PR 06-APR-1999; 99GB-0007805.
PR 15-APR-1999; 99GB-0008645.
PA (DANI-) DANISCO AS.
PI Sibbesen O, Sorensen JF.
WP1; 2000-465744/40.
PT Mutant xylanase protein identified using xylanase inhibitor useful for preparing non-sticky dough for bakery products -

XX Claim 24; Page 112; 112pp; English.
 PS
 CC The present sequence is derived from an endo-beta-1,4-xylanase
 CC inhibitor. The protein is obtained from wheat flour. The specification
 CC also describes a mutant xylanase protein. The xylanase is useful for
 CC preparing a foodstuff, preferably a bakery product or a substance
 CC (e.g. a dough) for making the bakery product. Wild type xylanase or
 CC mutant xylanase is useful for preparing a dough that is less sticky
 CC than a dough comprising a fungal xylanase. The xylanase inhibitor is
 CC useful for screening high degree resistance xylanases for dough
 CC preparation. The xylanase is also useful for preparing a non-sticky
 CC dough. A combination of xylanase and the inhibitor is useful for
 CC calibrating and/or determining the quantity of inhibitor in a wheat
 CC flour sample.
 CC
 SQ Sequence 31 AA:
 Query Match 100.0%; Score 145; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLPRGSTGVAGLANSGLALPAQVASAOK 31
 DB 1 LASLPRGSTGVAGLANSGLALPAQVASAOK 31
 RESULT 2
 ID AAE14690 standard; peptide; 31 AA.
 XX AAE14690;
 AC
 XX
 XX
 DT 21-AUG-2002 (first entry)
 DE Wheat flour xylanase inhibitor B chain Lys-C digested fragment #1.
 XX
 XX Refrigerated dough; syrruping; arabinoxylan; bakery product; bread;
 KW pizza base; cake; biscuit; wheat; flour; xylanase inhibitor.
 XX
 OS Triticum aestivum.
 XX
 PN WO200152657-A1.
 PD 26-JUL-2001.
 PF 17-JAN-2001; 2001WO-1B00168.
 XX
 PR 18-JAN-2000; 2000GB-0001136.
 PA (DANI-) DANISCO AS.
 XX
 PI Poulsen CH, Sorensen JF;
 DR WPI; 2001-457446/49.
 XX
 PT Production of refrigerated dough with reduced syrruping, useful in
 PT production of bakery products such as bread, comprises admixing cereal
 PT flour, water and protein that prevents enzymatic degradation of
 PT arabinoxylan in the cereal flour -
 PS
 PS Disclosure; Page 23; 26pp; English.
 CC The invention relates to a process for producing refrigerated dough
 CC with reduced 'syrruping' (precipitation of liquid on the dough surface
 CC because of a reduction in water holding capacity caused by the breakdown
 CC of arabinoxylan over time). The process comprises admixing cereal flour
 CC and water with a protein that reduces/prevents enzymatic degradation of
 CC arabinoxylan in the cereal flour. The preferred protein is a xylanase
 CC inhibitor. The method is useful to produce refrigerated dough in which
 CC syrruping is reduced or eliminated. Refrigerated dough is typically
 CC stored for long periods to enable fresh baked products (e.g. bread,
 CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
 CC requirements using the method by the use of specific proteins/protein
 CC combinations. The present sequence is wheat flour
 CC endo-beta-1,4-xylanase inhibitor B chain Lys-C digested fragment.
 XX
 SQ Sequence 31 AA:
 Query Match 100.0%; Score 145; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLPRGSTGVAGLANSGLALPAQVASAOK 31
 DB 1 LASLPRGSTGVAGLANSGLALPAQVASAOK 31
 RESULT 3
 ID AAU07395 standard; protein; 31 AA.
 XX AAU07395;
 AC
 XX
 XX
 DT 18-DEC-2001 (first entry)
 DE Bacillus subtilis xylanase inhibitor #4.
 XX
 XX Bacillus subtilis xylanase inhibitor #4.
 KW xylanase; plant cell wall; baking; cereal; starch production; wood;
 KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
 XX
 OS Bacillus subtilis.
 XX
 PN WO200166711-A1.
 PD 13-SEP-2001.
 PF 08-MAR-2001; 2001WO-1B00426.
 XX
 PR 08-MAR-2000; 2000GB-0005585.
 XX
 PR 27-JUN-2000; 2000GB-0015751.
 PA (DANI-) DANISCO AS.
 XX
 PI Sibbesen O, Sorensen JF;
 DR WPI; 2001-596834/67.
 XX
 PT Novel variant xylanase polypeptide or its fragment useful for degrading
 PT or modifying plant cell wall, comprises amino acid modifications such
 PT that the polypeptide has altered sensitivity to xylanase inhibitor -
 PS
 PS Disclosure; Page 63; 70pp; English.
 CC The invention relates to a variant xylanase polypeptide (I) or its
 CC fragment having xylanase activity, comprising one or more amino acid
 CC modifications such that (I) or its fragment has an altered sensitivity to
 CC a xylanase inhibitor as compared with the parent xylanase enzyme. (I) or
 CC its coding sequence (II) is useful for degrading or modifying plant cell
 CC wall or for processing a plant material by contacting the plant cell wall
 CC or plant material with (I) or (II). (I) is useful for modifying plant
 CC materials, and in baking, processing cereals, starch production,
 CC processing wood and enhancing the bleaching of wood pulp. (I)
 CC is useful for altering the viscosity derived from the presence of
 CC hemicellulose or arabinoxylan in a solution or system comprising plant
 CC cell wall material. (I) is useful for preparing a foodstuff such as
 CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
 CC The present sequence represents the amino acid sequence of Bacillus
 CC subtilis xylanase inhibitor #4 as described in the method of the
 CC invention.
 XX
 SQ Sequence 31 AA:
 Query Match 100.0%; Score 145; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;

Matches 31: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 LLASLPRGSTGVAGLSGLALPAQVASAOK 31
 |||
 Db 1 LLASLPRGSTGVAGLSGLALPAQVASAOK 31

RESULT 4
 AAU75815
 ID AAU75815 standard; Protein: 370 AA.
 AC AAU75815:
 XX
 DT 23-APR-2002 (first entry)
 XX

Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.
 DE
 XX Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 XX Triticum aestivum cultivar Soissons.
 OS
 XX WO200198474-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcours J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR N-PSDB; ABK13672.
 DR WPI: 2002-114579/15.
 DR
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).
 CC
 XX
 SQ Sequence 370 AA:
 XX

Query Match 100.0%; Score 145; DB 23; Length 370.
 Best Local Similarity 100.0%; Pred. No. 3.3e-12;
 Matches 31: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 LLASLPRGSTGVAGLSGLALPAQVASAOK 31
 |||
 Db 118 LLASLPRGSTGVAGLSGLALPAQVASAOK 148

RESULT 5
 AAU75816
 ID AAU75816 standard; Protein: 370 AA.
 AC AAU75816:
 XX
 DT 23-APR-2002 (first entry)
 XX

Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.02.
 DE
 XX Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 XX Triticum aestivum cultivar Estica.
 OS
 XX WO200198474-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcours J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR N-PSDB; ABK13673.
 DR WPI: 2002-114579/15.
 DR
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 57; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syripping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malling of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).

XX Sequence 370 AA:

Query Match 100.0%; Score 145; DB 23; Length 370;
 Best Local Similarity 100.0%; Pred. No. 3.3e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLASIPRGSTGVAGLANSGLAPAOVASAK 31
 ||||||||||||||||||||||||||||
 DB 118 LLASIPRGSTGVAGLANSGLAPAOVASAK 148

RESULT 6
 AAU75813
 ID AAU75813 standard; Protein; 381 AA.

XX AAU75813;

DT 23-APR-2002 (first entry)

XX Wheat L endoxylanase inhibitor, TAXI I, variant #2.

XX Wheat, TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KM immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
 KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KM noodle; animal feed; starch separation; maize processing; malling;
 KM plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX Triticum aestivum.
 OS

XX key Location/Qualifiers

FT Misc-difference 95 /label= Unknown

FT Misc-difference 98 /label= Unknown

FT Misc-difference 101 /label= Unknown

FT Misc-difference 110 /label= Unknown

FT Misc-difference 333 /label= Unknown

PN WO200198474-A1.

XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcourt J., Debysse W., Gebruers K., Goesaert H., Pierens K., Robben J.
 PI Van Campenhout S;
 DR WPI: 2002-114579/15.

XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 XX Claim 127; Page 51; 127pp; English.

XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syripping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malling of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

XX Sequence 381 AA:

Query Match 100.0%; Score 145; DB 23; Length 381;
 Best Local Similarity 100.0%; Pred. No. 3.4e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLASIPRGSTGVAGLANSGLAPAOVASAK 31
 ||||||||||||||||||||||||||||
 DB 129 LLASIPRGSTGVAGLANSGLAPAOVASAK 159

RESULT 7
 AAU75812

ID	AAU75812 standard; Protein; 381 AA.
XX	
AC	AAU75812;
XX	
DT	23-APR-2002 (first entry)
DE	Wheat L endoxylanase inhibitor, TAXI I, variant #1.
XX	
KW	Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW	xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW	immobilised enzyme; enzyme; dough syruiping; cereal product; beer;
KW	straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW	noodle; animal feed; starch separation; maize processing; malting;
KW	plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX	
OS	Triticum aestivum.
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 95 /label= Unknown
FT	Misc-difference 98 /label= Unknown
FT	Misc-difference 101 /label= Unknown
FT	Misc-difference 110 /label= Unknown
FT	Misc-difference 333 /label= Unknown
FT	Misc-difference /label= Unknown
XX	
PN	WO200198474-A1.
XX	
PD	27-DEC-2001.
XX	
PE	21-JUN-2001; 2001WO-BE00106.
XX	
PR	22-JUN-2000; 2000GB-0015296.
PR	25-JAN-2001; 2001GB-0002018.
PR	26-JAN-2001; 2001GB-0002194.
PR	16-MAR-2001; 2001GB-0006564.
PR	21-MAY-2001; 2001GB-0012328.
XX	
PA	(LEUV-) LEEVEN RES & DEV.
XX	
PI	DeJoucou J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
PI	Van Campenhout S;
XX	
DR	WPI: 2002-114579/15.
XX	
PT	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT	beta-glucanolytic enzymes comprises using endoxylanases during
PT	screening for inhibition activity or affinity chromatography with
PT	immobilised enzymes -
XX	
PS	Claim 127; Page 50-51; 127pp; English.
XX	
CC	The invention relates to separating and/or isolating inhibitors of
CC	cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
CC	screening the inhibition activity by using two or more enzymes during
CC	separation and/or isolation steps that allow to distinguish inhibitors of
CC	different specificity or by using an affinity chromatographic step with
CC	immobilised enzymes and/or antibodies against inhibitors. Also
CC	included are an isolated nucleic acid molecule encoding an inhibitor
CC	which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
CC	alpha-L-arabino-furanosidase and/or other cellulose, xylan,
CC	arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
CC	nucleic acid, an expression system transformed with the nucleic
CC	acid, a host organism transformed with the nucleic acid, the inhibitory
CC	proteins encoded by the nucleic acids and modulators of the proteins.
CC	A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC	plant or plant materials transformed with the nucleic acid are useful
CC	for the formation of an endoxylanase-inhibitor complex, screening
CC	endoxylanases that are totally, less or not inhibited by the inhibitors,
CC	reducing syruiping in refrigerated dough compositions, affecting the

CC	relative affinity and/or relative hydrolysis specificity and/or relative
CC	hydrolysis rate versus water-extractable and/or water-unextractable
CC	arabinoxyans of endoxylanases such as by the formation of an
CC	endoxylanase/inhibitor complex, improving the malting of cereals such as
CC	barley, sorghum and wheat and/or the production of beer, improving the
CC	production and/or quality of baked or extruded cereal products such as
CC	straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC	biscuits, pasta and noodles, animal feed stuff, improving the production
CC	of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten
CC	starch separation and production, improving maize processing,
CC	plant disease resistance and nutraceutical and/or pharmaceutical
CC	applications, improving paper and pulp technologies. The present
CC	sequence is a variant of wheat TAXI-I ('. aestivum L endoxylanase
CC	inhibitor).
CC	Note: Variant amino acids are highlighted in the specification but
CC	no wild-type sequence is shown for comparison.
CC	
XX	
SO	Sequence 381 AA:
	Query Match 96.6%; Score 140; DB 23; Length 381;
	Best Local Similarity 96.8%; Pred. No. 1.7e-11;
	Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY	1 LLASIPRGSTGVAGLANSGLALPAOVASAK 31
Db	129 LLASIPRGSTGVAGLADSLALPAOVASAK 159
RESULT 8	
AAU75814	AAU75814 standard; Protein: 381 AA.
XX	
AC	AAU75814;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.
XX	
KW	Wheat: TAXI-I, L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW	xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW	immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
KW	straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW	noodle; animal feed; starch separation; maize processing; malting;
KW	plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX	
OS	Triticum aestivum.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 95
FT	/label= Unknown
FT	Misc-difference 98
FT	/label= Unknown
FT	Misc-difference 101
FT	/label= Unknown
FT	Misc-difference 110
FT	/label= Unknown
FT	Misc-difference 145
FT	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference 183
FT	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference 232
FT	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference 275
FT	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference 282
FT	/label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
FT	Misc-difference 333
FT	/label= Unknown
XX	
PN	WO200198474-A1.
XX	
PD	27-DEC-2001.
XX	

PF 21-JUN-2001; 2001WO-BE00106.
 XX
 PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI; 2002-114579/15.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 51; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a microheterogenic variant of wheat TAXI-I (T. aestivum L
 CC endoxylanase inhibitor)
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 SQ Sequence 381 AA:
 Query Match 95.2%; Score 138; DB 23; Length 381;
 Best Local Similarity 96.8%; Pred. No. 3.2e-11;
 Matches 30; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LLASLPRGSTGVAAGLNSGLALPAOVASAK 31
 DB 129 LLASLPRGSTGVAAGLNSGLALPAOVASAK 159
 RESULT 9
 AAU75818
 ID AAU75818 standard; Protein: 277 AA.
 XX
 AC AAU75818;

XX
 DT 23-APR-2002 (first entry)
 XX
 DE L endoxylanase inhibitor, ATXI-II, partial sequence ATXI-II.01.
 XX
 KW ATXI-II: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor; ds;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Aegilops tauschii variety Acc2220051.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 4 /note= "Encoded by CTC"
 FT
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI; 2002-114579/15.
 DR N-PSDB: ABK13675.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 58; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC

CC	plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence represents a partial Aegilops tauschii ATXI-II (Aegilops tauschii L endoxylanase inhibitor).
XX	
SO	Sequence 277 AA;
OY	Query Match 86.2%; Score 125; DB 23; Length 277; Best Local Similarity 87.1%; Pred. No. 1.5e-09; Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
DB	1 LLASLPGSGTGAAGLANSLAPAOVASQK 31 118 LLASLPGSGMGVAGLGSGGLAPAOVASTRK 148
RESULT 10	
ID	AAU75819 standard; Protein: 287 AA.
AC	AAU75819;
DT	23-APR-2002 (first entry)
DE	Wheat L endoxylanase inhibitor, partial sequence TAXI-III.
KW	Wheat; TAXI-III; L endoxylanase inhibitor; cellulolytic enzyme inhibitor; xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor; immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant; straight dough; sponge dough; Chorleywood Bread; biscuits; pasta; noodle; animal feed; starch separation; maize processing; malting; plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
OS	Triticum aestivum cultivar Solissons.
PN	WO200198474-A1.
PD	27-DEC-2001.
PX	21-JUN-2001; 2001WO-BE00106.
PF	22-JUN-2000; 2000GB-0015296. 25-JAN-2001; 2001GB-0002018. 26-JAN-2001; 2001GB-0002194. 16-MAR-2001; 2001GB-0006564. 21-MAY-2001; 2001GB-0012328.
PX	(LEUV-) LEUVEN RES & DEV.
PI	Delcourt J, Debysser W, Gebruers K, Goesaert H, Fierens K, Robben J, Van Campenhout S;
DR	N-PSSD; ABK13676.
PT	MP: 2002-114579/15.
CC	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or beta-glucanolytic enzymes comprises using endoxylanases during screening for inhibition activity or affinity chromatography with immobilised enzymes -
PS	Claim 127; Page 58; 127pp; English.
CC	The invention relates to separating and/or isolating inhibitors of cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising separating the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulose, xylan, arabinoxylian or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic

[illegible]

PI Van Camphenout S;
 XX WPI; 2002-114579/15.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 XX Claim 127; Page 63; 127pp; English.
 XX
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syneruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of barley HVXI-I (H. vulgare L endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 CC
 XX
 XX Sequence 185 AA;
 SO

Query Match 53.8%; Score 78; DB 23; Length 185;
 Best Local Similarity 89.5%; Pred. No. 0.0038;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 13 AGLANSGALPQVVASAK 31
 DB 1 AGFAGSGALPQVVASAK 19

RESULT 12
 AAU75825
 ID AAU75825 standard; Protein: 185 AA.
 AC AAU75825;
 XX
 XX 23-APR-2002 (first entry)
 DT
 XX
 XX 3arley L endoxylanase inhibitor, HVXI I, variant #2.
 DE
 XX
 XX Barley; HVXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syneruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX
 OS Hordeum vulgare.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 72
 FT Misc-difference 185
 FT Misc-difference 185
 FT Misc-difference 185
 XX /label= Unknown
 XX /label= Unknown
 PN W0200198474-A1.
 XX
 XX 27-DEC-2001.
 PD
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0005654.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEUVEN RES & DEV.
 PA
 XX
 XX Delcour J, Debysse W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Camphenout S;
 XX
 XX WPI; 2002-114579/15.
 DR
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 PT
 XX
 XX Claim 127; Page 63; 127pp; English.
 XX
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syneruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of barley HVXI-I (H. vulgare L endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 CC
 XX
 XX Sequence 185 AA;
 SO

Query Match 53.8%; Score 78; DB 23; Length 185;

Best Local Similarity 89.5%; Pred. No. 0.0038;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 13 AGLANGSLALPAQVASAOK 31
||| ||||| ||||| |||||
Db 1 AGFAGSGLALPAQVASAOK 19

RESULT 13

AAU75826
ID AAU75826 standard; Protein; 185 AA.

AC AAU75826;

DT 23-APR-2002 (first entry)

DE Barley L endoxylanase inhibitor, HVXI I, microheterogenic variant.

XX Barley; HVXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;

KW xylooligolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;

KW immobilised enzyme; enzyme; dough syruing; cereal product; beer; plant;

KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;

KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Hordeum vulgare.

XX Synthetic.

XX Key

FT Misc-difference 72 Location/Qualifiers

FT Misc-difference 128 /Label= Unknown

FT Misc-difference 132 /Label= Pro, Ala, Asp, Leu, Gly, Asn, Cys

FT Misc-difference 134 /Label= Pro, Ala, Asp, Leu, Gly, Asn, Cys

FT Misc-difference 148 /Label= Pro, Ala, Asp, Leu, Gly, Asn, Cys

FT Misc-difference 185 /Label= Pro, Ala, Asp, Leu, Gly, Asn, Cys

FT Misc-difference 185 /Label= Unknown

XX WO200198474-A1.

XX 27-DEC-2001.

XX 21-JUN-2001; 2001WO-BE00106.

XX 22-JUN-2000; 2000GB-0015296.

XX 25-JAN-2001; 2001GB-0002018.

XX 16-MAR-2001; 2001GB-0002194.

XX 21-MAY-2001; 2001GB-0012328.

XX (LEUV-) LEUVEN RES & DEV.

XX Delcourt J, Debysier W, Gebuwers K, Goesaert H, Fierens K, Robben J;

XX Van Campenhout S;

XX WPI; 2002-114579/15.

XX Separating and/or isolating inhibitors of cellulolytic, xylooligolytic, or

XX beta-glucanolytic enzymes comprises using endoxylanases during

XX screening for inhibition activity or affinity chromatography with

CC immobilised enzymes and/or antibodies against inhibitors. Also

CC included are an isolated nucleic acid molecule encoding an inhibitor

CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,

CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,

CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the

CC nucleic acid, a host organism transformed with the nucleic acid, the inhibitory

CC proteins encoded by the nucleic acids and modulators of the proteins.

CC A recombinant protein, glycoprotein or polypeptide or microorganisms,

CC plant or plant materials transformed with the nucleic acid are useful

CC for the formation of an endoxylanase-inhibitor complex, screening

CC endoxylanases that are totally, less or not inhibited by the inhibitors,

CC reducing syruing in refrigerated dough compositions, affecting the

CC relative affinity and/or relative hydrolysis specificity and/or relative

CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an

CC endoxylanase/inhibitor complex, improving the malting of cereals such as

CC barley, sorghum and wheat and/or the production of beer, improving the

CC production and/or quality of baked or extruded cereal products such as

CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,

CC biscuits, pasta and noodles, animal feed stuff, improving the production

CC of starch derived syrups, sorbitol, xylitol, wheat gluten

CC starch separation and production, improving maize processing,

CC plant disease resistance and nutraceutical and/or pharmaceutical

CC applications, improving paper and pulp technologies. The present

CC sequence is a microheterogenic variant of barley HVXI-I (H. vulgare L

CC endoxylanase inhibitor).

CC Note: Variant amino acids are highlighted in the specification but

CC no wild-type sequence is shown for comparison.

XX SQ Sequence 185 AA;

Query Match 53.8%; Score 78; DB 23; Length 185;

Best Local Similarity 89.5%; Pred. No. 0.0038;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 13 AGLANGSLALPAQVASAOK 31

Db 1 AGFAGSGLALPAQVASAOK 19

XX ||| ||||| ||||| |||||

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 52999.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137228.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151990.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	04-MAY-1999	9905-0132484
PR	05-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132486
PR	06-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-0132488
PR	11-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-0134221
PR	14-MAY-1999	9905-0134370
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0135424
PR	20-MAY-1999	9905-0135441
PR	21-MAY-1999	9905-0135553
PR	24-MAY-1999	9905-0135629
PR	25-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136392
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139452
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139899
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140554
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141842
PR	01-JUL-1999	9905-0141842
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142927
PR	12-JUL-1999	9905-0143277
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143524
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144321
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145088
PR	21-JUL-1999	9905-0145088

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.1%; Score 64; DB 21; Length 433;
Best Local Similarity 41.4%; Pred. NO. 0.89; Mismatches 10; Indels 0; Gaps 0;
Matches 12; Conservative 7;

Oy 1 LLASIPRGSTGVAGIANSGLAPQVASA 29
Db 165 LKGLAKSTVGMAGMRNIGLPSQFAA 193

Search completed: May 20, 2003, 16:14:41
Job time : 34.7253 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen, Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 12.6044 Seconds
(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-16

Sequence: 1 LLASLPKRGSTGVAGLANSGLALPAQVASAOK 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	46.9	433	2	T14329
2	66	45.5	427	1	S06750
3	64	44.1	433	2	F86163
4	62	42.8	434	2	G86163
5	55	37.9	456	1	TVHDMC
6	55	37.9	460	2	S11558
7	53	37.9	464	1	TVHDM2
8	53	36.6	330	2	T29640
9	53	36.6	520	2	H64510
10	52	35.9	513	2	T40998
11	51	35.2	406	2	A53010
12	50.5	34.8	503	2	G75262
13	50	34.5	468	2	T34679
14	50	34.5	1114	2	T18197
15	49.5	34.1	310	2	JN0490
16	49.5	34.1	672	2	H83315
17	49.5	34.1	1443	2	G75393
18	49.5	34.1	1752	2	A45407
19	49	33.8	135	2	D64339
20	49	33.8	377	2	C91270
21	49	33.8	377	2	C91270
22	49	33.8	652	1	S39409
23	48.5	33.4	157	2	AC2775
24	48.5	33.4	162	2	E97555
25	48.5	33.4	485	2	G65051
26	48.5	33.4	485	2	B85920
27	48.5	33.4	485	2	C91075
28	48.5	33.4	1222	2	G59100
29	48	33.1	189	2	AC3420

30	48	33.1	244	2	F70834	hypothetical prote
31	48	33.1	473	2	C87325	nitrate transpor
32	48	33.1	601	1	B56564	transcription fac
33	48	33.1	608	2	T03476	conserved hypoth
34	48	33.1	613	2	T00077	gag-like protein
35	48	33.1	719	2	T00266	hypothetical prote
36	48	33.1	750	2	H70711	probable mult prot
37	48	33.1	839	2	F75518	hypothetical prote
38	48	33.1	1046	2	F71432	hypothetical prote
39	47.5	32.8	238	1	LRR1B2	clathrin light cha
40	47.5	32.8	446	2	T50786	nucleoid DNA-bind
41	47.5	32.8	3788	2	T30851	lysosomal traffick
42	47	32.4	72	2	A11026	hypothetical prote
43	47	32.4	216	2	I51234	paired box transcr
44	47	32.4	324	2	H83127	probable dehydrog
45	47	32.4	337	2	F83877	acetoin dehydrogen

ALIGNMENTS

RESULT 1
T14329
dermal glycoprotein precursor, extracellular - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Nov-2000
C:Accession: T14329
R:Salch, S.; Sturm, A.; Fujii, T.; Chrispeels, M.J.
Planta 188, 432-438, 1992
A>Title: CDNA cloning of an extracellular dermal glycoprotein of carrot and its expe
A:Reference number: Z17987
A:Accession: T14329
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-433 <SAT>
A:Cross-references: EMBL:D14550; NID:d1004922; PID:d1003923
A:Experimental source: subspecies US-Harumakigosun; root
C:Genetics:
A:Gene: EDGp1
C:Superfamily: conglutin gamma
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-433/Product: dermal glycoprotein, extracellular #status predicted <MAT>

Query Match 46.9%; Score 68; DB 2; Length 433;
Best Local Similarity 51.7%; Pred. No. 0.21;
Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 1 LLASLPKRGSTGVAGLANSGLALPAQVASA 29
DB 163 LLONLASGVGMAGLGRTRIALPSQFASA 191

RESULT 2
S06750
basic 7S globulin precursor - soybean
N:Alternate names: insulin-binding protein
C:Species: Glycine max (soybean)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S06750; S00218; A61003; JN0311; JN0312
R:Kagawa, H.; Hirano, H.
Nucleic Acids Res. 17, 8868, 1989
A>Title: Sequence of a cDNA encoding soybean basic 7S globulin.
A:Reference number: S06750; NID:90067863; PMID:2587227
A:Accession: S06750
A:Molecule type: mRNA
A:Residues: 1-427 <KAL>
A:Cross-references: EMBL:X16469; NID:g18542; PID:CA34489.1; PID:g18543
R:Kagawa, H.; Yamauchi, F.; Hirano, H.
FEBS Lett. 226, 145-149, 1987
A>Title: Soybean basic 7S globulin represents a protein widely distributed in legume
A:Reference number: S00218
A:Accession: S00218
A:Molecule type: protein

A:Residues: 276-293,'X','295-304','EX','423-427','X','1VPGPGCLPNONGVTSLGPMMXMPAR','QL','356, 42','AX' <KAG>
 R:Hirano, H.; Watanabe, T.
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 A:Title: Microsequencing of proteins electrotransferred onto immobilizing matrices from
 A:Reference number: A61003; MUID:91031409; PMID:2226413
 A:Accession: A61003
 A:Molecule type: protein
 A:Residues: 25-57;135-150;276-304,'X','306-308','X','310-311;383-417 <HIR>
 R:Barbashov, S.F.; Egorov, T.A.; Kochkina, V.M.
 Biotorg. Khim. 17, 421-423, 1991
 A:Title: Isolation and characterization of soybean insulin-binding protein.
 A:Reference number: JN0311; MUID:91291181; PMID:2064630
 A:Accession: JN0311
 A:Molecule type: protein
 A:Residues: 'VPIPHKRN','27-33','X','35-42','XX','45','X','47','X','49','AX','52 <BAR>
 A:Note: 20-Thr, 21-Asn, 25-Pro and 50-Asn were also found
 A:Accession: JN0312
 A:Molecule type: protein
 A:Residues: 'S','99','IV','102','SX','105','C','286-299','XX','302','XX','305 <BA2>
 A:Note: 102-Gly and 296-Lys were also found
 C:Comment: This protein specifically binds to bovine insulin.
 C:Superfamily: conglutin gamma
 C:Keywords: heterodimer; seed; storage protein
 F:1-4/Domain: signal sequence #status predicted <SIG>
 F:25-275/Product: basic 7S globulin light chain #status experimental <MAT1>
 F:275-427/Product: basic 7S globulin light chain #status experimental <MAT2>

Query Match 45.5% Score 66; DB 1: Length 427;
 Best Local Similarity 54.2% Pred. No. 0.37;
 Matches 13: Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 LPRGSGVAGLANSGALPAAQVSA 28
 DB 175 LPRNGVAGLGHAPISLPNOLAS 198

RESULT 3
 F66163
 Hypothetical protein F15K9.17 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 A:Accession: F66163
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzila,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Molecule type: DNA
 A:Accession: F66163
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-433 <STO>
 A:Cross-references: GB:AE005172; NID:g3850579; PIDN:AC72119.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: conglutin gamma

Query Match 44.1% Score 64; DB 2: Length 433;
 Best Local Similarity 41.4% Pred. No. 0.7;
 Matches 12: Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 1 LLASLPRGSGVAGLANSGALPAAQVSA 29
 DB 165 LKGLAKGVAGMAGMRHNIIGLPSPFAA 193

RESULT 4

G66163
 Hypothetical protein F15K9.16 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 A:Accession: G66163
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzila,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Accession: G66163
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 <STO>
 A:Cross-references: GB:AE005172; NID:g3850580; PIDN:AC72120.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: conglutin gamma

Query Match 42.8% Score 62; DB 2: Length 434;
 Best Local Similarity 41.4% Pred. No. 1.3;
 Matches 12: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 1 LLASLPRGSGVAGLANSGALPAAQVSA 29
 DB 166 LKGLAKGVAGMAGMRHNIIGLPSPFAA 194

RESULT 5
 TVHUMC
 Transforming protein N-myc (version 1) - human
 C:Species: Homo sapiens (man)
 C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 15-Oct-1999
 A:Accession: A01355; 159514; S02249
 R:Slamon, L.W.; Schwab, M.; Bishop, J.M.
 Proc. Natl. Acad. Sci. U.S.A. 83, 1772-1776, 1986
 A:Title: Nucleotide sequence of the human N-myc gene.
 A:Reference number: A01355; MUID:86149381; PMID:2869488
 A:Accession: A01355
 A:Molecule type: DNA
 A:Residues: 1-456 <STA>
 A:Cross-references: GB:M13241; NID:g189247; PIDN:AAA36371.1; PID:g386983
 R:Slamon, D.J.; Boone, T.C.; Seeger, R.C.; Keith, D.E.; Chazin, V.; Lee, H.C.; Souza,
 Science 232, 768-772, 1986
 A:Title: Identification and characterization of the protein encoded by the human N-myc
 A:Reference number: 159514; MUID:86179901; PMID:3008339
 A:Accession: 159514
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 63-456 <RES>
 A:Cross-references: GB:M13228; NID:g189243; PIDN:AAA36370.1; PID:g386982
 R:Ibson, J.M.; Rabbits, P.H.
 Oncogene 2, 399-402, 1988
 A:Title: Sequence of a germ-line N-myc gene and amplification as a mechanism of activ
 A:Reference number: S00945; MUID:88202932; PMID:2834684
 A:Accession: S02249
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-354,'V','356-456 <IBS>
 A:Cross-references: EMBL:Y00664; NID:g35074; PIDN:CAA68678.1; PID:g35076
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1988
 C:Genetics:
 A:Gene: GDB:MYCN; NMNC
 A:Cross-references: GDB:120527; OMIM:164840
 A:Map position: 2p24.3-2p24.3
 A:Introns: 256/1
 C:Superfamily: myc transforming protein; myc transforming protein homology

hypothetical protein MJEC108 - Methanococcus jannaschii plasmid pMB800
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 29-Sep-1999
 C:Accession: H64510
 R:Ball, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 Isori, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kane, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: H64510
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-520 <BUIL>
 A:Cross-references: GB:L77118; NID:g1500644; TIGR:MJEC108; PIDN:AAC37081.1; PID:g1522648
 C:Genetics:
 A:Map position: ECLREV7391-5829
 A:Genome: plasmid
 A:Start codon: GTG
 A:Note: this stable 58-kilobase pair plasmid is also designated ECL (large extrachromosc
 C:Superfamily: conserved hypothetical protein MJ1429

Query Match 36.6%; Score 53; DB 2; Length 520;
 Best Local Similarity 53.8%; Pred. No. 24;
 Matches 14; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 1 LASLPRGSTGVAGLANGLALPAQV 26
 DB 475 LLQDLPLQSTGEATVGVSSLPPLALV 500

RESULT 10
 T40998
 hypothetical protein SPOC1450.16c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T40998
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Voicakart, G.
 submitted to the EMBL Data Library, March 1999
 A:Reference number: Z21962
 A:Accession: T40998
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-513 <LYN>
 A:Cross-references: EMBL:AL049559; PIDN:CAB40183.1; GSPDB:GN00068; SPDB:SPOC1450.16c
 A:Experimental source: strain 972h.; cosmid c1450
 C:Genetics:
 A:Gene: SPDB:SPOC1450.16c
 A:Map position: 3

Query Match 35.9%; Score 52; DB 2; Length 513;
 Best Local Similarity 45.5%; Pred. No. 32;
 Matches 15; Conservative 2; Mismatches 6; Indels 10; Gaps 1;

OY 1 LASLPRGSTGV-----AGLANGLALP 23
 DB 339 LLAKLPDGTSTVCTPKNTFWPYAGLPNTGRSNP 371

RESULT 11
 AS3010
 copper transport protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: Protein P9642.3; protein YPR124W
 C:Species: Saccharomyces cerevisiae
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
 C:Accession: AS3010; S69016; S48876
 R:Danis, A.; Yuan, D.S.; Hallie, D.; Askwith, C.; Elide, D.; Mochle, C.; Kaplan, J.; Kian
 Cell 76, 393-402, 1994
 A:Title: Molecular characterization of a copper transport protein in Saccharomyces cerev
 A:Reference number: AS3010; MUID:94123345; PMID:8293472
 A:Accession: AS3010
 A:Molecule type: DNA

A:Residues: 1-406 <DAN>
 A:Cross-references: GB:U02511; NID:g407947; PIDN:AAAI7369.1; PID:g407948
 R:Johnson, D.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of S. cerevisiae cosmid 9642.
 A:Reference number: S69014
 A:Accession: S69016
 A:Molecule type: DNA
 A:Residues: 1-406 <JOH>
 A:Cross-references: EMBL:U40828; NID:g1066467; PIDN:ABB68064.1; PID:g1066470; MIPS:YP
 C:Genetics:
 A:Gene: SGD:CYP1
 A:Cross-references: SGD:S0006328; MIPS:YPR124W
 A:Map position: 16R
 A:Function:
 A:Description: copper transport
 C:Keywords: copper transport; transmembrane protein
 F:156-172/Domain: transmembrane #status predicted <TM1>
 F:255-271/Domain: transmembrane #status predicted <TM2>

Query Match 35.2%; Score 51; DB 2; Length 406;
 Best Local Similarity 37.9%; Pred. No. 34;
 Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

OY 2 LASLPRGSTGVAGLANGLALPAQVASAO 30
 DB 67 MASMSGSSMSGMSMSSTPTSSASAO 95

RESULT 12
 G75262
 hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75262
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
 ; M.; Shen, M.; Vamathevan, J.D.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75262
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-503 <WHT>
 A:Cross-references: GB:AE002082; GB:AE000513; NID:g6460347; PIDN:AAFI2072.1; PID:g646
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2528
 A:Map position: 1

Query Match 34.8%; Score 50.5; DB 2; Length 503;
 Best Local Similarity 42.9%; Pred. No. 50;
 Matches 12; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANGLAL--PAOVASA 29
 DB 81 IPAGMGITAGLLSTGAAOTSPAPVTAA 108

RESULT 13
 T34679
 probable protoporpyrinogen oxidase - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Nov-2000
 C:Accession: T34679
 R:Harits, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21553
 A:Accession: T34679
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-468 <HAR>

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 8.34615 Seconds

(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-16

Sequence: 1 LLASLPKSGTGVAGLANSLALPAQVSAQK 31

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	45.5	427	1	7SBG_SOYBN
2	55	37.9	460	1	P13917 glycine max
3	55	37.9	464	1	MYCN_MARMO
4	53	36.6	312	1	MYCN_HUMAN
5	53	36.6	520	1	DIF1_CAEEL
6	51	35.2	406	1	YZ08_METJA
7	49.5	34.1	742	1	CTRL1_YEAST
8	49	33.8	124	1	TBX3_HUMAN
9	49	33.8	652	1	SPEH_METJA
10	48.5	33.4	485	1	NOS2_PARDE
11	48	33.1	601	1	PTDA_ECOLI
12	48	33.1	750	1	PDML_DROME
13	47.5	32.8	229	1	MUTB_MICTU
14	47.5	32.8	1509	1	CLCB_RAT
15	47	32.4	416	1	GSRI_HUMAN
16	47	32.4	216	1	PAX6_CHICK
17	47	32.4	422	1	PAX6_COTJA
18	47	32.4	422	1	PAX6_MOUSE
19	47	32.4	422	1	PAX6_HUMAN
20	47	32.4	649	1	PAX6_XENLA
21	47	32.4	1324	1	PTMA_VIBCH
22	46.5	32.1	162	1	IRS2_HUMAN
23	46.5	32.1	419	1	PUR6_BACSU
24	46.5	32.1	1088	1	CARP_CRYPA
25	46	31.7	208	1	PIGO_HUMAN
26	46	31.7	396	1	FCP_ISOGA
27	46	31.7	442	1	FCP_ARATH
28	46	31.7	442	1	AROA_BORBR
29	46	31.7	748	1	AROA_BORPE
30	46	31.7	748	1	PTIP_ECOLI
31	46	31.7	952	1	PTIP_SALTY
32	46	31.7	1267	1	UVRA_THIEH
33	46	31.7	1267	1	UVRA_REOVD
					VL3_REOVD

34	46	31.7	1267	1	VL3_REOVL
35	46	31.7	1434	1	VG65_HSV11
36	45.5	31.4	456	1	RINI_RAT
37	45.5	31.4	3801	1	LYST_HUMAN
38	45	31.0	111	1	RLA2_ASPEU
39	45	31.0	323	1	JUND_CHICK
40	45	31.0	341	1	YR15_MYCTU
41	45	31.0	381	1	AMPC_ENTCL
42	45	31.0	578	1	VCSI_BACSU
43	45	31.0	730	1	EIS_HUMAN
44	45	31.0	2505	1	CCAA_HUMAN
45	45	31.0	4568	1	DYHB_CHLRE

ALIGNMENTS

RESULT 1
7SBG_SOYBN STANDARD: PRT: 427 AA.
AC P13917: Q39901: Q43464:
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic 7S globulin precursor (Bg) (Sbg7S).
GN BG OR G7S.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Miyagishirume; TISSUE=Seed;
RX MEDLINE=90067863; PubMed=2587227;
RA Kagawa H., Hirano H.;
RT "Sequence of a cDNA encoding soybean basic 7S globulin.";
RL Nucleic Acids Res. 17:8868-8868(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Miyagishirume; TISSUE=Etolated leaf;
RX MEDLINE=94336768; PubMed=8058830;
RA Watanabe Y., Hirano H.;
RT "Nucleotide sequence of the basic 7S globulin gene from soybean.";
RL Plant Physiol. 105:1019-1020(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Shi-shi; TISSUE=Cotyledon;
RA Shu T.F., Hsieh K.L., Hsing Y.I., Chen Z.Y., Chow T.Y.;
RT Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE.
RA Kagawa H., Yamauchi F., Hirano H.;
RT "Soybean basic 7S globulin represents a protein widely distributed in legume species.";
RL FEBS Lett. 226:145-149(1987).
CC - FUNCTION: SEED STORAGE PROTEIN. HAS A PROTEIN KINASE ACTIVITY.
CC BINDS LEGNINOLIN.
CC - FUNCTION: SEED STORAGE PROTEIN.
CC - SUBUNIT: THE MATURE BG IS COMPOSED OF HIGH- AND LOW-KDa SUBUNITS WHICH ARE LINKED BY DISULFIDE BONDING.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: X16469; CAA34489.1; -
CC EMBL: U59425; AAB0390.1; -
CC EMBL: D16107; BAA03681.1; -
DR

DR PIR: S06750; S06750.
 KW Glycoprotein; Seed storage protein; signal.
 FT SIGNAL 1 24
 FT CHAIN 25 427 BASIC 7S GLOBULIN.
 FT CHAIN 25 427 HIGH-KDA SUBUNIT.
 FT CHAIN 276 427 LOW-KDA SUBUNIT.
 FT CONFLICT 48 48 W -> S (IN REF. 3).
 FT CONFLICT 264 264 N -> T (IN REF. 1).
 FT CONFLICT 305 305 F -> C (IN REF. 1).
 FT CONFLICT 309 309 F -> C (IN REF. 1).
 SQ SEQUENCE 427 AA; 46393 MW; 66041BC0680BACB CRC64;

Query Match 45.5%; Score 66; DB 1; Length 427;
 Best Local Similarity 54.2%; Pred. No. 0.15;
 Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 LPRSTGVAGLANSGLALPAOVAS 28
 DB 175 LPRMTGVAGLGHAPISLPNOLAS 198
 ||| : ||||| : : ||| : |||

RESULT 2
 MYCN_MARMO STANDARD; PRT: 460 AA.
 ID MYCN_MARMO
 AC Q61976;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE N-myc proto-oncogene protein (N-myc1).
 GN MYCN OR NMYC OR MYC1.
 OS Marmota monax (Woodchuck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
 OC Marmota.
 CC NCBI_TaxID=9995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90370481; PubMed=2395655;
 RA Jorel G., Tjollals P., Buendia M.-A.;
 RL Nucleotide sequence of the woodchuck N-myc gene (WN-myc1).";
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
 CC
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR IMBL: X53673; CAA37712.1; -
 DR IMBL: X53674; CAA37712.1; JOINED.
 DR HSSP: P25912; IHLO.
 DR TRANSFAC: T03440; -
 DR InterPro: IPR001092; HLH_basic.
 DR InterPro: IPR002418; TF_Myc.
 DR Pfam: PF00010; HLH_1.
 DR Pfam: PF01056; Myc_N_term; 1.
 DR PRINTS: PRO0044; LEUZIPPMYC.
 DR SMART: SM00353; HLH_1.
 DR PROSITE: PS00038; HLH_1; 1.
 DR PROSITE: PS50888; HLH_2; 1.
 KW Nuclear protein; DNA-binding; proto-oncogene; Phosphorylation.
 FT LNA_BIND 258 274 ASP/GLU-RICH (ACIDIC).
 FT LNA_BIND 377 390 BASIC DOMAIN.
 FT LNA_BIND 391 430 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT LNA_BIND 429 450 LEUCINE-ZIPPER (POTENTIAL).
 FT MOD_RES 257 257 PHOSPHORYLATION (BY CK2)

FT FT (BY SIMILARITY).
 FT MOD_RES 259 259 PHOSPHORYLATION (BY CK2)
 FT FT (BY SIMILARITY).
 SQ SEQUENCE 460 AA; 49192 MW; 8A16686C82F5B02E CRC64;
 SO SEQUENCE 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Query Match 37.9%; Score 55; DB 1; Length 460;
 Best Local Similarity 52.4%; Pred. No. 4.7;
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 7 RSTGVAGLANSGLALPAOVA 27
 DB 159 RGHGTAGAGAGALPAELA 179
 ||| : ||| : : ||| : |||

RESULT 3
 MYCN_HUMAN STANDARD; PRT: 464 AA.
 ID MYCN_HUMAN
 AC P04198;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE N-myc proto-oncogene protein.
 GN MYCN OR NMYC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66149381; PubMed=2869488;
 RA Stanton L.W., Schwab M., Bishop J.M.;
 RL Nucleotide sequence of the human N-myc gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:1772-1776(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86092232; PubMed=3510398;
 RA Kohl N.E., Legouy E., Depinho R.A., Nisen P.D., Smith R.K.,
 RA Gee C.E., Alt F.W.;
 RL "Human N-myc is closely related in organization and nucleotide
 RL sequence to c-myc.";
 RL Nature 319:73-77(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88202932; PubMed=2834684;
 RA Ibsen J.M., Rablitts P.H.;
 RL "Sequence of a germ-line N-myc gene and amplification as a mechanism
 RL of activation.";
 RL Oncogene 2:399-402(1988).
 RN [4]
 RP SEQUENCE OF 71-464 FROM N.A.
 RX MEDLINE=86179901; PubMed=3008339;
 RA Slamon D.J., Boone T.C., Seeger R.C., Keith D.E., Chazin V., Lee H.C.,
 RA Souza L.M.;
 RL "Identification and characterization of the protein encoded by the
 RL human N-myc oncogene.";
 RL Science 232:768-772(1986).
 RN [5]
 RP SEQUENCE OF 313-464 FROM N.A.
 RX MEDLINE=85215633; PubMed=2987858;
 RA Michitsch R.W., Melera P.W.;
 RL "Nucleotide sequence of the 3' exon of the human N-myc gene.";
 RL Nucleic Acids Res. 13:2545-2558(1985).
 RN [6]
 RP PHOSPHORYLATION BY CK2.
 RX MEDLINE=93049352; PubMed=1425701;
 RA Hagihara T., Nakaya K., Nakamura Y., Nakajima H., Nishimura S.,
 RA Taya Y.;
 RL "Specific phosphorylation of the acidic central region of the N-myc
 RL protein by casein kinase II.";
 RL Eur. J. Biochem. 209:945-950(1992).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.


```

CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING FETAL DEVELOPMENT.
CC -1- DISEASE: AMPLIFICATION OF THE N-MYC GENE IS ASSOCIATED WITH A
CC VARIETY OF HUMAN TUMORS, MOST FREQUENTLY NEUROBLASTOMA, WHERE THE
CC LEVEL OF AMPLIFICATION APPEARS TO INCREASE AS THE TUMOR
CC PROGRESSES.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/NMYC112.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X03294; CAA27037.1; -
CC DR EMBL: X03295; CAA27038.1; -
CC DR EMBL: M13241; AAA36371.1; ALT_INIT.
CC DR EMBL: M13228; AAA36370.1; -
CC DR EMBL: X02363; -; NOT_ANNOTATED_CDS.
CC DR EMBL: Y00664; CAA68678.1; ALT_INIT.
CC DR PIR: A25744; TVHUM2.
CC DR PIR: A01355; TVHUMC.
CC DR PIR: A22937; A22937.
CC DR PIR: S02249; S02249.
CC DR HSSP: P25912; IHLO.
CC DR TRANSFAC: T02379; -
CC DR Genem: HGNC:7559; MYCN.
CC DR MIM: 164840; -
CC DR InterPro: IPR001092; HLH_basic.
CC DR InterPro: IPR002418; TF_Myc.
CC DR Pfam: PF00010; HLH_1.
CC DR Pfam: PF01056; MYC_N_term; 1.
CC DR PRINTS: PR00044; LEUZIPPRMVC.
CC DR SMART: SM00353; HLH_1.
CC DR PROSITE: PS00038; HLH_1; 1.
CC DR PROSITE: PS50888; HLH_2; 1.
CC KW Nuclear protein; DNA-binding; Proto-oncogene; Phosphorylation.
CC FT DOMAIN 262 278 ASP/GLU-RICH (ACIDIC).
CC FT DNA_BIND 381 394 BASIC DOMAIN.
CC FT DOMAIN 395 434 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
CC FT DOMAIN 433 454 LEUCINE-ZIPPER (POTENTIAL).
CC FT MOD_RES 261 261 PHOSPHORYLATION (BY CK2).
CC FT MOD_RES 263 263 PHOSPHORYLATION (BY CK2).
CC FT CONFLICT 227 227 A -> P (IN REF. 2).
CC FT CONFLICT 363 363 I -> V (IN REF. 3).
CC SQ SEQUENCE 464 AA; 49561 MW; 560E885602E30DAD CRC64;

Query Match 37.9%; Score 55; DB 1; Length 464;
Best Local Similarity 52.4%; Pred. No. 4.8;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2.
RX MEDLINE-95292971; PubMed-7774589;
RA Ahinger J.;
RT "Embryonic tissue differentiation in Caenorhabditis elegans requires
RL dif-1, a gene homologous to mitochondrial solute carriers.";
RN EMBL J. 14:2307-2316(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95102108; PubMed-7803813;
RA Runswick M.J., Philippides A., Lauria G., Walker J.E.;
RT "Extension of the mitochondrial transporter super-family: sequences of
RL five members from the nematode worm, Caenorhabditis elegans.";
RN DNA Seq. 4:281-291(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Murray J., Wohlmann P.;
RN Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP REVISIONS.
RA Waterston R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Seems to play a role in the maintenance of tissue
CC differentiation in the developing embryo, but not for its
CC initiation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (potential).
CC -1- DEVELOPMENTAL STAGE: Although dif-1 activity is required only
CC during embryogenesis, it is expressed at all stages of
CC development.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z48240; CAA88283.1; -
CC DR EMBL: X76115; CAA53721.1; -
CC DR EMBL: U61949; AAB03153.2; -
CC DR Wormep: F49E8.5; CE28408.
CC DR InterPro: IPR001993; Mitoch_carrier.
CC DR Pfam: PF00153; mito_carrier; 3.
CC DR PROSITE: PS00215; MITOCH_CARRIER; 3.
CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC Developmental protein.
CC FT TRANSMEM 5 25 POTENTIAL.
CC FT TRANSMEM 69 89 POTENTIAL.
CC FT TRANSMEM 104 124 POTENTIAL.
CC FT TRANSMEM 209 229 POTENTIAL.
CC FT TRANSMEM 277 297 POTENTIAL.
CC SQ SEQUENCE 312 AA; 33134 MW; CCF1300CA4D64F CRC64;

Query Match 36.6%; Score 53; DB 1; Length 312;
Best Local Similarity 38.7%; Pred. No. 6;
Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

```

```

RESULT 4
DIF1_CAEEL STANDARD; PRT; 312 AA.
AC Q27257; Q20630; (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Protein dif-1.
DE DIF-1 OR F49E8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.

```

```

RESULT 5
Y208_METJA STANDARD; PRT; 520 AA.
AC Q60270;
DT 01-NOV-1997 (rel. 35, Created)

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJEC108.
 GN MJEC108.
 OS Methanococcus jannaschii.
 OC Archaea: Euryarchaeota: Methanococci: Methanococcales:
 OC Methanocaldococcaceae: Methanocaldococcus.
 NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Butt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uitterlbeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii."
 RL Science 273:1058-1073(1996).
 CC -I- SIMILARITY: TO M.JANNASCHII MJ1429 AND MJ1565.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: L77118; AAC37081.1;
 DR TIGR: MJEC108;
 DR InterPro: IPR002789; DUF87.
 DR Pfam: PF01935; DUF87; 1
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 520 AA: 58180 MW; D1888AFDFE1D9079 CRC64;
 Query Match 36.6%; Score 53; DB 1; Length 520;
 Best Local Similarity 53.8%; Pred No. 9.8; Mismatches 10; Indels 0; Gaps 0;
 Matches 14; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 QY 1 LLASLPKSTGVAGLANSGLAPAOV 26
 DB 475 LLQDLPLQSTGEAIVGVSSILPALV 500
 RESULT 6
 CTR1_YEAST STANDARD: PRT: 406 AA.
 ID CTR1_YEAST
 AC P49573;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Copper transport protein CTR1 (Copper transporter 1).
 GN CTR1 OR YPR124W OR P9642.3.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94123345; PubMed=8293472;
 RA Dancis A., Yuan D.S., Haile D., Askwith C., Eide D., Moehle C.,
 RA Kaplan J., Klausner R.D.;
 RT "Molecular characterization of a copper transport protein in S.
 RT cerevisiae: an unexpected role for copper in iron transport.";
 RL Cell 76:393-402(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288c / AB972;

RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
 RA Araujo R., Aparicio A., Bartell B.G., Badcock K., Benes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., Dipolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hedling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunkeler-Smith S., Hyman R., Johnston M., Kaiman S., Klehe K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mitalpali S., Moestl D.,
 RA Mueller-Auer S., Namath A., Neuwirth U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
 RA Scherrens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
 RA Uristetaru L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambolt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Han J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=95014372; PubMed=7929270;
 RA Dancis A., Haile D., Yuan D.S., Klausner R.D.;
 RT "The Saccharomyces cerevisiae copper transport protein (Ctr1p).
 RT Biochemical characterization, regulation by copper, and physiologic
 RT role in copper uptake.";
 RL J. Biol. Chem. 269:25660-25667(1994).
 CC -I- FUNCTION: REQUIRED FOR HIGH AFFINITY COPPER (PROBABLY REDUCED CU
 CC -I- TRANSPORT INTO THE CELL.
 CC -I- SUBUNIT: OLIGOMER.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- INDUCTION: BY COPPER DEPRIVATION, AND REPRRESSED BY COPPER
 CC SUFFICIENCY.
 CC -I- PTM: O-GLYCOSYLATED EXTENSIVELY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: U02511; AAA17369.1;
 DR EMBL: U40828; AAB80864.1;
 DR SGD: S0006328; CTR1.
 DE Transmembrane; Transport; Copper; Glycoprotein; Repeat.
 GN TRANSMEMBRANE; TRANSPORT; MET/SER-RICH.
 FT DOMAIN 1 127 3 X 19 AA TANDEN REPEATS OF S-M-X-M-X-A-
 FT FT 9 65 M-S-S-A-S-K-T-X-X-S-X-M-X.
 FT REPEAT 9 27 1.
 FT REPEAT 28 46 2.
 FT REPEAT 47 65 3.
 FT DOMAIN 114 117 POLY-SER.
 FT TRANSMEM 153 173 POTENTIAL.
 FT TRANSMEM 251 271 POTENTIAL.
 FT DOMAIN 272 406 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 406 AA: 44441 MW; 569C9CB19C1117E8 CRC64;
 Query Match 35.2%; Score 51; DB 1; Length 406;
 Best Local Similarity 37.9%; Pred No. 14;
 Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
 QY 2 LLASLPKSTGVAGLANSGLAPAOVASQ 30
 DB 67 KASMSMGSSMSGMSMSSTPTSSASQ 95
 RESULT 7
 TBX3_HUMAN
 ID TBX3_HUMAN STANDARD: PRT: 742 AA.

AC 015119;09KRF8;
DT 15-JUL-1999 (Rel.38, Created)
DT 16-OCT-2001 (Rel.40, Last sequence update)
DT 15-JUN-2002 (Rel.41, Last annotation update)
DE T-box transcription factor TBX3 (T-box protein 3).
GN TBX3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM I), AND FUNCTION.
RC TISSUE-Breast carcinoma;
RX MEDLINE=99398688; PubMed=10465586;
RA He M.-L., Wen L., Campbell C.E., Wu Y.Y., Rao Y.;
RT "Transcription repression by Xenopus Et and its human ortholog TBX3, a
RT gene involved in ulnar-mammary syndrome.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10212-10217(1999).
RN [2]
RP SEQUENCE OF 1-488 FROM N.A. (ISOFORM I).
RC TISSUE-Kidney;
RX MEDLINE=97351519; PubMed=9207801;
RA Banshad M., Le T., Watkins W.S., Dixon M.E., Kramer B.E., Roeder A.D.,
RA Carey J.C., Root S., Schinzel A., Van Maldergem L., Gardner R.J.M.,
RA Lin R.C., Seidman C.E., Seidman J.G., Wallerstein R., Moran E.,
RA Sutphen R., Campbell C.E., Jorde L.B.;
RT "The spectrum of mutations in TBX3: genotype/phenotype relationship in
RT ulnar-mammary syndrome.";
RL Nat. Genet. 16:311-315(1997).
RN [3]
RP SEQUENCE OF 591-742 FROM N.A. (ALTERNATIVE SPLICING, AND VARIANTS UMS.
RX MEDLINE=99264236; PubMed=10330342;
RA Banshad M., Le T., Watkins W.S., Dixon M.E., Kramer B.E., Roeder A.D.,
RA Carey J.C., Root S., Schinzel A., Van Maldergem L., Gardner R.J.M.,
RA Lin R.C., Seidman C.E., Seidman J.G., Wallerstein R., Moran E.,
RA Sutphen R., Campbell C.E., Jorde L.B.;
RT "The spectrum of mutations in TBX3: genotype/phenotype relationship in
RT ulnar-mammary syndrome.";
RL Am. J. Hum. Genet. 64:1550-1562(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM II).
RC TISSUE-Adrenal gland;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL
CC -I- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN DEVELOPMENTAL
CC PROCESSES. PROBABLY PLAYS A ROLE IN LIMB PATTERN FORMATION.
CC -I- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -I- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; I, II (SHOWN HERE) AND
CC III, ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORMS II AND III
CC CONTAIN AN INTERRUPTED T-BOX DOMAIN. AN ADDITIONAL ISOFORM IV MAY
CC BE PRODUCED BY JOINING EXON 1 TO EXON 7 THEREBY ELIMINATING THE T-
CC BOX.
CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -I- DISEASE: DEFECTS IN TBX3 ARE THE CAUSE OF ULNAR-MAMMARY SYNDROME
CC (UMS). THIS DISEASE IS CHARACTERIZED BY ULNAR RAY DEFECTS,
CC OBESITY, HYPOGENITALISM, DELAYED PUBERTY, HYPOPLASIA OF NIPPLES
CC AND APOCRINE GLANDS.
CC -I- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF170708; AAD50989.2; -
CC DR EMBL; AF002228; AAC12947.1; -
CC DR EMBL; AF140240; AAC1816.1; -
CC DR EMBL; AF216750; AAF61207.1; -
CC HSSP; P24781.1YBR.
CC TRANSFAC; T04413; -.

DR	TRANSFAC: T04A14; -
DR	Gemew; HGNC: 11602; TBX3.
DR	MIM; 601621; -
DR	InterPro; IPR001699; TF_T-box.
DR	Pfam; PF00907; T-box; 1.
DR	PRINTS; PR00937; TBOX.
DR	SMART; SM00425; TBOX; 1.
DR	PROSITE; PS01283; TBOX_1; 1.
DR	PROSITE; PS01264; TBOX_2; 1.
DR	PROSITE; PSS0252; TBOX_3; 1.
KW	Transcription regulation; DNA-binding; Repressor; Nuclear protein;
KW	Developmental protein; Disease mutation; Alternative splicing.
FT	DNA_BIND 112 220
FT	T_BOX (FIRST PART).
FT	T_BOX (SECOND PART).
FT	TRANSCRIPTION REPRESSION DOMAIN.
FT	MISSING (IN ISOFORM I).
FT	AAHLAGCPJPGICFPAGIAGOEFNGHPILFLPISOFAMGCA
FT	FSSMAAGKGPLLATIVSGASTVGSGLDSTANASAAAQIGLS
FT	GASAATLPFHLDQHVASOGIAMSFGSLFFPYTYMAAAAP
FT	AASLRPOLRCTAPLL -> RSSVHRHFPR (IN
FT	ISOFORM III).
FT	MISSING (IN ISOFORM II).
FT	L -> P (IN UMS).
FT	/FTId=VAR_009601.
FT	Y -> S (IN UMS).
FT	/FTId=VAR_009602.
FT	K -> Q (IN REF. 4).
FT	LQPOLRCTAPL -> SAASSVHRHF (IN REF.
FT	3).
FT	A -> V (IN REF. 3).
FT	L -> P (IN REF. 1).
SQ	SEQUENCE 742 AA; 79402 MW; D2178A2480962160 CRC64;
Query Match	34.1%; Score 49.5; DB 1; Length 742;
Besht Local Similarity	43.3%; Pred. No. 40;
Matches 13; Conservative 7; Mismatches 7; Indels 3; Gaps	
Qy	1 LLASLPRGSTGVAGLIANSGLPAQVAQAQ 30 : : : :
Db	542 LLATVSGASTVGSLDSTAM--ASAAQAQ 568
RESULT 8	
ID	SPEH_METJA STANDARD: PRT: 124 AA.
AC	057763;
DT	01-NOV-1997 (Rel. 35, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	S-deoxylymethionine decarboxylase preenzyme (EC 4.1.1.50) (AdomercPC)
DE	(Samdc) [Contains: S-deoxylymethionine decarboxylase beta chain; S-
DE	adenosylmethionine decarboxylase alpha chain].
CN	SPEH OR MC0315.
OC	Methanococcus jannaschii.
OC	Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC	Methanocaldococcaceae; Methanocaldococcus.
OX	NCBI_TaxID=2190;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA	MEDLINE=96337999; PubMed=8688087;
RA	Built C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Goeysne J.D.,
RA	Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA	Scott T.L., Geohagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA	Uuterkack T.R., Kelley J.M., Peterson J.D., Sadov P.W., Hanna M.C.,
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA	Klenk H.-P., Fraser C.M., Smith H.O., Woesie C.R., Venter J.C.,
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus
JL	jannaschii".
Science 273:	1058-1073(1996).

```

RN [2]
RP CHARACTERIZATION.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-20528316; PubMed-11073910;
RA Kim A.D., Graham D.E., Seeholzer S.H., Markham G.D.;
RT "5-adenosylmethionine decarboxylase from the archaeon Methanococcus
  jannaschii: identification of a novel family of pyruvoyl enzymes.";
RL J. Bacteriol. 182:6667-6672(2000).
CC -1- FUNCTION: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
  AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE BIOSYNTHESIS FROM
  PUTRESCINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5'-deoxy-5'-
  adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC -1- COFACTOR: Pyruvoyl group.
CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
CC -1- MASS SPECTROMETRY: MM=6794.5; METHOD=MALDI; RANGE=1-63.
CC -1- MASS SPECTROMETRY: MM=6991.6; METHOD=MALDI; RANGE=64-124.
CC -1- MISCELLANEOUS: THERMOSTABLE.
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC ADOMETDC FAMILY. SUBFAMILY
  1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67486; AAB98301.1; ALT_INIT.
DR TIGR: MJ0315; -.
DR InterPro: IPR003826; SAMDC.
DR Pfam: PF02675; DUF206; 1.
KW Spermidine biosynthesis; Lyase; Decarboxylase; zymogen; Pyruvate;
  Complete proteome.
FT CHAIN 1 63 S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
FT FT CHAIN.
FT FT CHAIN 64 124 S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
FT FT CHAIN.
FT SITE 63 64 CLEAVAGE (NONHYDROLYTIC).
FT MOD_RES 64 64 CONVERTED TO A PYRUVOL GROUP.
FT ACT_SITE 84 84 IMPORTANT FOR CATALYTIC ACTIVITY (BY
  SIMILARITY).
FT FT SEQUENCE 124 AA; 13792 MW; BC1152CDE080E113 CRC64;
  SO
  Query Match 33.8%; Score 49; DB 1; Length 124;
  Best Local Similarity 58.8%; Pred. No. 8.2;
  Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
  Oy 6 PRGSTGVAGLANSGLAL 22
  1:1111111111111111
  Db 52 POGATGVAVLAESHIAI 68
  RESULT 9
  NOSZ_PARDE STANDARD: PRT; 652 AA.
  ID NOSZ_PARDE STANDARD: PRT; 652 AA.
  AC 051705;
  DT 01-NOV-1997 (Rel. 35, Created)
  DT 01-NOV-1997 (Rel. 35, Last sequence update)
  DT 01-NOV-1997 (Rel. 35, Last annotation update)
  DE Nitrous-oxide reductase precursor (EC 1.7.99.6) (N(2)OR).
  GN NOSZ.
  OS Paracoccus denitrificans.
  OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
  OC Bacteroidetes;
  OX NCBI_TaxID=266;
  RN [11]
  RP SEQUENCE FROM N.A.
  RC STRAIN=NLB8944;
  RX MEDLINE-94062841; PubMed-8243476;
  RA Hoeren F.U., Berks B.C., Ferguson S.J., McCarthy J.E.G.;
  RT "Sequence and expression of the gene encoding the respiratory

```

```

RT nitrous-oxide reductase from Paracoccus denitrificans. New and
  conserved structural and regulatory motifs."
RL Eur. J. Biochem. 218:49-57(1993).
CC -1- FUNCTION: NITROUS-OXIDE REDUCTASE IS PART OF A BACTERIAL
  RESPIRATORY SYSTEM WHICH IS ACTIVATED UNDER ANAEROBIC
  CONDITIONS IN THE PRESENCE OF NITRATE OR NITROUS OXIDE.
CC -1- CATALYTIC ACTIVITY: N(2) + H(2)O + acceptor = nitrous oxide +
  reduced acceptor.
CC -1- COFACTOR: COPPER (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER; EACH SUBUNIT CONTAINS 2 DINUCLEAR COPPER
  CENTERS A AND 2. 2 IS THOUGHT TO BE THE SITE OF NITROUS OXIDE
  REDUCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, TO MITOCHONDRIAL AND
  BACTERIAL COX2 SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74792; CA52798.1; -.
KW Signal; Oxidoreductase; Copper; Metal-binding; Periplasmic.
FT SIGNAL 1 57 BY SIMILARITY.
FT CHAIN 1 57 NITROUS-OXIDE REDUCTASE.
FT METAL 58 652 COPPER A 1 (BY SIMILARITY).
FT METAL 595 595 COPPER A 1 AND 2 (BY SIMILARITY).
FT METAL 630 630 COPPER A 2 (BY SIMILARITY).
FT METAL 632 632 COPPER A 2 (BY SIMILARITY).
FT METAL 634 634 COPPER A 1 AND 2 (BY SIMILARITY).
FT METAL 638 638 COPPER A 2 (BY SIMILARITY).
FT METAL 641 641 COPPER A 1 (BY SIMILARITY).
FT METAL 641 641 COPPER A 1 (BY SIMILARITY).
SQ SEQUENCE 652 AA; 71413 MW; 40492A4FDE7EDEAB CRC64;
  Query Match 33.8%; Score 49; DB 1; Length 652;
  Best Local Similarity 54.5%; Pred. No. 41;
  Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
  Oy 3 ASLPRGSTGVAGLANSGLALPA 24
  1:1111111111111111
  Db 37 AALGLGTAGVAVAGSAAALAA 58
  RESULT 10
  PTDA_ECOLI STANDARD: PRT; 485 AA.
  ID PTDA_ECOLI STANDARD: PRT; 485 AA.
  AC P24241; Q46880;
  DT 01-MAR-1992 (Rel. 21, Created)
  DT 01-NOV-1997 (Rel. 35, Last sequence update)
  DT 15-JUN-2002 (Rel. 41, Last annotation update)
  DE PTS system, arbutin-, cellobiose-, and salicin-specific IIABC
  component (EIIBC-ASC) (Arbutin-, cellobiose-, and salicin-permease
  IIABC component) (Phosphotransferase enzyme II, ABC component)
  DE (EC 2.7.1.69) (EIIBC-ASC).
  GN ASCF OR B2713.
  OS Escherichia coli.
  OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
  OC Escherichia.
  OX NCBI_TaxID=562;
  RN [11]
  RP SEQUENCE FROM N.A.
  RC STRAIN=K12;
  RX MEDLINE-92334140; PubMed-1630307;
  RA Hall B.G., Xu L.;
  RT "Nucleotide sequence, function, activation, and evolution of the
  cryptic asc operon of Escherichia coli K12.";
  RL Mol. Biol. Evol. 9:688-706(1992).
  RN [12]
  RP SEQUENCE FROM N.A.
  RC STRAIN=K12 / MG1655;
  RX MEDLINE-97426617; PubMed-9278503;

```

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HRP). IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR.
CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIIB DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M73326; AAA16429.1; -
CC EMBL: U29579; AAA69225.1; -
CC EMBL: AE000355; AAC75757.1; -
CC PIR: B44070; B44070.
CC PIR: S27552; S27552.
CC HSSP: P05053; 1IBA.
CC Ecogene: EGI0086; ascf.
CC InterPro: IPR004719; PTSIIC_glic.
CC InterPro: IPR001996; PTS_EIIB.
CC InterPro: IPR003352; PTS_EIIC.
CC Pfam: PF00367; PTS_EIIB.1.
CC Pfam: PF02378; PTS_EIIC.1.
CC Pfam: PD001476; PTS_EIIB.1.
CC TIGRfams: TIGR00826; EIIB.1.
CC TIGRfams: TIGR00852; pts-glic.1.
CC PROSITE: PS01035; PTS_EIIB_CYS.1.
CC KW Phosphotransferase system; Sugar transport; Transferase;
CC Phosphorylation; Transmembrane; Inner membrane; Complete proteome.
CC FT DOMAIN 1 43 EIIB DOMAIN.
CC FT DOMAIN ? ? EIIB DOMAIN.
CC FT MOD_RES ? 485 EIIC DOMAIN.
CC FT MOD_RES 28 28 PHOSPHORYLATION (BY SIMILARITY).
CC FT TRANSMMEM 102 122 PHOSPHORYLATION (BY SIMILARITY).
CC FT TRANSMMEM 146 166 POTENTIAL.
CC FT TRANSMMEM 177 197 POTENTIAL.
CC FT TRANSMMEM 207 227 POTENTIAL.
CC FT TRANSMMEM 254 274 POTENTIAL.
CC FT TRANSMMEM 285 305 POTENTIAL.
CC FT TRANSMMEM 330 350 POTENTIAL.
CC FT TRANSMMEM 363 383 POTENTIAL.
CC FT TRANSMMEM 389 409 POTENTIAL.
CC FT TRANSMMEM 433 453 POTENTIAL.
CC FT CONFLICT 167 170 HLPR -> OSAA (IN REF. 1).
CC FT CONFLICT 311 311 R -> H (IN REF. 1).
CC SEQUENCE 485 AA; 51229 MW; CFEEDDD8CABC838 C6C64;
CC
CC Query Match 33.4%; Score 48.5; DB 1; Length 485;
CC Best Local Similarity 40.0%; Pred. No. 36;
CC Matches 14; Conservative 4; Mismatches 12; Indels 5; Gaps 1;
CC
CC 1 LIASLPRG-----STGVAGLANSGALPAQVSAQ 30
CC DB 389 LIASLSGFTCGAVAGMAGLASHMAAPGLFTSVQ 423

RESULT 11
PDM1_DROME
ID PDM1_DROME STANDARD: PRT: 601 AA.
AC P31368;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nubbin protein (Twain protein) (POU domain protein 1) (PDM-1) (DPOU-
DE 19) (DPOU1).
GN NUB OR TWN OR PDM-1 OR POU-19 OR OCT1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92001544; PubMed=1680380;
RA Billin A.N., Cockerill K.A., Poole S.J.;
RT "Isolation of a family of Drosophila POU domain genes expressed in
RT early development.";
RL Mech. Dev. 34:75-84(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9214419; PubMed=1685891;
RA Lloyd A., Sakonju S.;
RT "Characterization of two Drosophila POU domain genes, related to
RT oct-1 and oct-2, and the regulation of their expression patterns.";
RL Mech. Dev. 36:87-102(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91352045; PubMed=1881906;
RA Dick T., Yang X., Yeo S., Chia W.;
RT "Two closely linked Drosophila POU domain genes are expressed in
RT neuroblasts and sensory elements.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7645-7649(1991).
CC -1- FUNCTION: DNA-BINDING REGULATORY PROTEIN IMPLICATED IN EARLY
CC DEVELOPMENT. INVOLVED IN NEURONAL CELL FATE DECISION. REPRESSED
CC DIRECTLY OR INDIRECTLY BY THE BX-C HOMEOBOX PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: NEUROBLASTS AND SENSORY ELEMENTS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED PRIMARILY DURING THE FIRST HALF OF
CC EMBRYOGENESIS. INITIAL EXPRESSION IN CELLULAR BLASTODERM STAGE,
CC THEN IN ECTODERMAL STRIPES DURING GERMBAND EXTENSION. BROAD
CC EXPRESSION IN THE NEUROECTODERM FOLLOWED BY LIMITATION TO DISCRETE
CC SUBSETS OF CNS CELLS, AND EXPRESSION IN SPECIFIC PNS NEURONS AND
CC SUPPORT CELLS.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEOBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M81957; AAA28829.1; -
CC EMBL: S80561; AAR21409.1; -
CC EMBL: M65015; AAA28480.1; -
CC PIR: A41277; A41277.
CC TRANSFAC: T01900; -
CC Flybase: FBgn0002970; nub.
CC InterPro: IPR001356; Homeobox.
CC InterPro: IPR00327; POU_domain.
CC Pfam: PF00046; homeobox.1.
CC Pfam: PF00157; pou.1.
CC PRINTS: PR00028; POUDOMAIN.

```

DR ProDom: PD000010; Homeobox; 1.
DR ProDom: PD000583; POU domain; 1.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00352; POU; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00035; POU_1; 1.
DR PROSITE: PS00465; POU_2; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein;
KW Developmental protein;
FT DOMAIN 104 107 POLY-HIS.
FT DOMAIN 191 201 POLY-ALA.
FT DOMAIN 359 364 POLY-HIS.
FT DOMAIN 407 415 POLY-ALA.
FT DOMAIN 425 495 POU.
FT DNA_BIND 523 582 HOMEBOX.
FT CONFLICT 420 420 P->R (IN REF. 2).
SQ SEQUENCE 601 AA; 65202 MW; 8FCD7B3C162D2B2 CRC64;

Query Match 33.1%; Score 48; DB 1; Length 601;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 10 TGVAGLANSGLALPAQVASAQ 30
Db 272 SGLSLADDDPALTAQVAAQ 292

RESULT 12
MUTB_MYCTU STANDARD: PRT; 750 AA.
AC P71774;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable methylmalonyl-CoA mutase large subunit (EC 5.4.99.2) (MCM).
GN MUTB OR RV1493 OR MT1540 OR MTCY277.15.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Jordan S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badoock K., Besham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Jleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Jetermann J., Debey R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE ISOMERIZATION OF SUCCINYL-CoA TO
METHYLMALONYL-CoA DURING SYNTHESIS OF PROPIONATE FROM
TRICARBOXYLIC ACID-CYCLE INTERMEDIATES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: (R)-2-methyl-3-oxopropanoyl-CoA -> succinyl-
CoA.
CC -1- COFACTOR: ADENOSYLCOBALAMIN (BY SIMILARITY).
```

```

CC -1- PATHWAY: Propionic acid fermentation.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE METHYLMALONYL-CoA MUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z79701; CAB02043.1; -
DR EMBL: AE007022; AK45807.1; -
DR HSSP: P11653; IREQ.
DR TIGR: MT1540; -
DR Tuberculist; RV1493; -
DR InterPro: IPR003312; B12-binding.
DR InterPro: IPR001629; MM_COA_mutase.
DR Pfam: PF01642; MM_COA_mutase; 1.
DR Pfam: PF02310; B12-binding; 1.
DR TIGRFAMS: TIGR00640; acid_COA_mut_C; 1.
DR TIGRFAMS: TIGR00641; acid_COA_mut_N; 1.
DR PROSITE: PS00544; METHYLMALONYL-CoA_MUTASE; 1.
KM Isomerase; Vitamin B12; Cobalt; Complete proteome.
FT DOMAIN 587 626 METHYLMALONYL-CoA-BINDING SITE (BY
FT SIMILARITY).
FT METAL 629 629 COBALAMIN-BINDING (POTENTIAL).
FT FT 629 629 COBALT (POTENTIAL).
SQ SEQUENCE 750 AA; 80604 MW; 9E6B86CE507D022 CRC64;
```

```

QY 3 ASLPGSTGVAGLANSGLALPAQVASAQ 30
Db 534 AAEGGRAGADGNNLALALDARAQ 561

RESULT 13
CLCB_RAT STANDARD: PRT; 229 AA.
AC P08082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Clathrin light chain B (Lcb).
GN CLTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87178007; PubMed=3563513;
RA Kirchhausen T., Scarlato P., Harrison S.C., Monroe J.J., Chow E.P.,
RA Mattaliano R.J., Ramachandran K.L., Smart J.E., Ahn A.H., Brosius J.;
RT "Clathrin light chains LCA and LCB are similar, polymorphic, and
RT share repeated heptad motifs."
RL Science 236:320-324(1987).
CC -1- FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYHEDRAL COAT OF
CC COATED PITS AND VESICLES.
CC -1- SUBUNIT: CLATHRIN COATS ARE FORMED FROM MOLECULES CONTAINING 3
CC HEAVY CHAINS AND 3 LIGHT CHAINS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND
CC VESICLES.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, BRAIN (SHOWN HERE) AND NON-
CC BRAIN; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 38.3242 Seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-16
Perfect score: 145
Sequence: 1 LASLPRGSTGVAGLANGLALPAQVASAOK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121	83.4	419	10	OBS1V0 oryza sativ
2	94	64.8	424	10	OBS1V1 oryza sativ
3	93	64.1	422	10	OBS1V3 oryza sativ
4	88	60.7	434	10	OBS1V5 oryza sativ
5	84	57.9	434	10	OBS1U4 oryza sativ
6	84	57.9	442	10	OBS1U6 oryza sativ
7	68	46.9	433	10	OBS1U9 oryza sativ
8	64	44.1	433	10	OBS1U4 oryza sativ
9	64	44.1	433	10	OBS1U4 oryza sativ
10	63	43.4	369	10	OBS1U9 oryza sativ
11	63	43.4	402	10	OBS1U9 oryza sativ
12	62	42.8	434	10	OBS1U9 oryza sativ
13	59	40.7	386	10	OBS1U9 oryza sativ
14	57.5	39.7	443	10	OBS1U7 oryza sativ
15	53.5	36.9	411	16	OBS1U7 oryza sativ
16	53	36.6	416	16	OBS1U7 oryza sativ

17	52	35.9	513	3	O9Y7P3 schizosacch
18	51	35.2	231	2	O93NE1 myxococcus
19	51	35.2	825	2	O9Z6O5 rhodobacter
20	50.5	34.8	503	16	O9YRG4 deinococcus
21	50.5	34.8	1093	3	O8XOR0 neuropeptide
22	50	34.5	583	10	O946Y6 chlamydomon
23	49.5	34.1	23	10	O560O8 streptomyces
24	49.5	34.1	653	10	O8W5H2 oryza sativ
25	49.5	34.1	672	16	O91OK3 pseudomonas
26	49.5	34.1	743	4	O8TB20 homo sapien
27	49.5	34.1	1443	16	O9RUC7 streptococcus
28	49.5	34.1	1752	5	O07265 strongyloce
29	49	33.8	210	2	O9RNU2 streptococc
30	49	33.8	217	1	O8XZ76 halobacteri
31	49	33.8	290	5	O9YTS1 drosophila
32	49	33.8	377	2	O9X5C0 escherichia
33	49	33.8	377	2	O9WVU6 escherichia
34	49	33.8	377	16	O8XDQ2 escherichia
35	49	33.8	653	16	O9RDB5 streptomyces
36	49	33.8	1010	2	O9XCV5 cellulosoma
37	49	33.8	2237	5	O9YI22 drosophila
38	48.5	33.4	162	16	O8UEY4 agrobacteri
39	48.5	33.4	166	16	O9CN27 pasteurilla
40	48.5	33.4	485	16	O8X842 escherichia
41	48.5	33.4	1222	2	O9X349 bacillus an
42	48	33.1	189	16	O8YGI7 brucella me
43	48	33.1	208	2	O87372 acetobacter
44	48	33.1	244	16	O53721 mycobacteri
45	48	33.1	279	16	O987A8 rhizobium 1

ALIGNMENTS

RESULT 1					
OBS1V0	PRELIMINARY:	PRT:	419 AA.		
AC OBS1V0:					
DT 01-JUN-2002 (TREMREL. 21, Created)					
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)					
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)					
DE Putative dermal glycoprotein.					
GN P0504E02.9.					
OS- Oryza sativa (japonica cultivar-group).					
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;					
OC Ehrhartoideae; Oryzaeae; Oryza.					
OX NCBI_TaxID=39947;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=CV. NIPPONBARE.					
RA Sasaki T., Matsumoto T., Yamamoto K.;					
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC					
RT clone:P0504E02."					
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.					
DR EMBL: AP003269; BAB89708.1; - 792BE5FAF3F3A8 CRC64;					
SQ SEQUENCE 419 AA; 43786 MW; 792BE5FAF3F3A8 CRC64;					
Query Match	83.4%;	Score 121;	DB 10;	Length 419;	
Best local similarity	86.7%;	Pred. No. 3;	1e-08;		
Matches 26;	Conservative	2;	Mismatches 2;	Indels 0;	Gaps 0;
OY 1 LASLPRGSTGVAGLANGLALPAQVASAOK 30					
DB 164 LALRLPRGATGVAGLANGLALPAQVASAOK 193					
RESULT 2					
OBS1V1	PRELIMINARY:	PRT:	424 AA.		
AC OBS1V1:					
DT 01-JUN-2002 (TREMREL. 21, Created)					
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)					

```

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative dermal glycoprotein.
GN P0504E02.8
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89707.1; -.
SQ SEQUENCE 424 AA; 44620 MW; 0DACD5CAC8852C34 CRC64;

Query Match
Best Local Similarity 64.8%; Score 94; DB 10; Length 424;
Matches 21; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 LLASLRPGSTGVAGLANSGLALPAOVASAK 31
Db 163 LLAKLPAGAVGAGLARTLALQAVARSOK 193

RESULT 3
OY 08S1V3 PRELIMINARY; PRT; 422 AA.
ID 08S1V3;
AC 08S1V3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative dermal glycoprotein.
GN P0504E02.6.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89705.1; -.
SQ SEQUENCE 422 AA; 44745 MW; 08CFE26346400A9F CRC64;

Query Match
Best Local Similarity 64.1%; Score 93; DB 10; Length 422;
Matches 21; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 LLASLRPGSTGVAGLANSGLALPAOVASAK 30
Db 165 LLASLPAGAVGAGLGRSLALHAQVAATO 194

RESULT 4
OY 08S1V5 PRELIMINARY; PRT; 434 AA.
ID 08S1V5;
AC 08S1V5;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative dermal glycoprotein.
GN P0504E02.3.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.

```

```

OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89703.1; -.
SQ SEQUENCE 434 AA; 45342 MW; 7F2549288A6B28B1 CRC64;

Query Match
Best Local Similarity 60.7%; Score 88; DB 10; Length 434;
Matches 18; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 1 LLASLRPGSTGVAGLANSGLALPAOVASAK 31
Db 166 LLGSLPGAVGAGLGSAPLSLPSOVASLTK 196

RESULT 5
OY 08S1U4 PRELIMINARY; PRT; 434 AA.
ID 08S1U4;
AC 08S1U4;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE P0504E02.16 protein.
GN P0504E02.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89714.1; -.
SQ SEQUENCE 434 AA; 46236 MW; 87C36211DD57FE72 CRC64;

Query Match
Best Local Similarity 57.9%; Score 84; DB 10; Length 434;
Matches 17; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 LLASLRPGSTGVAGLANSGLALPAOVAS 28
Db 167 LLASLPAGATGVAGFSRRLPSLSQLAA 194

RESULT 6
OY 08S1U6 PRELIMINARY; PRT; 442 AA.
ID 08S1U6;
AC 08S1U6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE P0504E02.14 protein.
GN P0504E02.14.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0504E02."

```

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003269; BAB89712.1; -
 SO SEQUENCE 442 AA; 47390 MW; 4EAAFA84B0CB8E6 CRC64;

Query Match 57.9%; Score 84; DB 10; Length 442;
 Best Local Similarity 60.7%; Pred. No. 0.0033;
 Matches 17; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 LLASUPRGSTGVAGLANSGLALPAOVASA 28
 DB 168 LLASUPGATGVAGFSRRPLSLPSOLA 195

RESULT 7

ID 005929 PRELIMINARY; PRT: 433 AA.
 AC 005929;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EDGP precursor (Fragment).
 GN EDGP1.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RA Satoh S.;
 RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RA Satoh S.;
 RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.

RT "cDNA cloning of an extracellular dermal glycoprotein of carrot and
 its expression in response to wounding.";
 RL Planta 188:432-438(1992).
 DR EMBL: D14550; BAA03413.1; -

KW Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 20 POTENTIAL.
 FT CHAIN 21 433 EDGP.
 SO SEQUENCE 433 AA; 45841 MW; 06EDFC19B7B8C38E CRC64;

Query Match 46.9%; Score 68; DB 10; Length 433;
 Best Local Similarity 51.7%; Pred. No. 0.47;
 Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 1 LLASUPRGSTGVAGLANSGLALPAOVASA 29
 DB 163 LLONTASGVAGMAGLGRTRIALPSOPASA 191

RESULT 8

ID 09ZVS4 PRELIMINARY; PRT: 433 AA.
 AC 09ZVS4;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F15K9.17 protein (Hypothetical 45.7 kDa protein) (Putative
 extracellular dermal glycoprotein EDGP).
 GN F15K9.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;

RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
 RA Kremenetskaia I., Luros J., Araujo R., Buehler E., Conway A.B.,
 RA Dwyer K., Feng J., Kim C., Li Y., Shin P., Sun H., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

[3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

[4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

[5]
 RN SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Theologis A., Ecker J., Davis R.W.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

[6]
 RN SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseima E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shin P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.17 (GI:3850579).";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

[7]
 RN SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseima E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shin P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.17 (GI:3850579).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC005278; AAC72119.1; -
 DR EMBL: AF325092; AAK17160.1; -
 DR EMBL: AY035098; AAL15204.1; -
 DR EMBL: AY035026; AAK59531.1; -
 KW Hypothetical protein.
 SO SEQUENCE 433 AA; 45717 MW; 7214FC4B8BA72962 CRC64;

Query Match 44.1%; Score 64; DB 10; Length 433;
 Best Local Similarity 41.4%; Pred. No. 1.6;
 Matches 12; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 1 LLASUPRGSTGVAGLANSGLALPAOVASA 29
 DB 165 LKGLAKGTVMAGMGRHNTGLPSOPAA 193

RESULT 9

ID 08RVH5 PRELIMINARY; PRT: 433 AA.
 AC 08RVH5;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE - Basic 7S globulin isoform precursor.

OS glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. MIYAGISHIROME; TISSUE=DEVELOPING SEEDS;
 RA Ishizu Y., Sassa H., Hirano H.;
 RT "Sequence of a cDNA encoding soybean basic 7S globulin isoform."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB084260; BAB91077.1; -;
 KW Signal.
 FT SIGNAL
 SO SEQUENCE 433 AA; 47205 MW; FA41B93DBD3EA38 CRC64;
 Query Match 44.1%; Score 64; DB 10; Length 433;
 Best Local Similarity 50.0%; Pred. No. 1.6;
 Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
 Oy 1 LLASIPRGSTGAGLANGLALPAQVAS 28
 Db 178 LQKGLPRNIGCVAGLGHAPISLPNOLAS 205
 RESULT 10
 09FSZ9 PRELIMINARY: PRT; 369 AA.
 ID 09FSZ9
 AC 09FSZ9
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE Putative extracellular dermal glycoprotein (Fragmant).
 OS Cicier arletum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
 RA Dopico B., Sanchez M.A., Labrador E.;
 RT "A putative extracellular dermal glycoprotein is expressed in chickpea epicotyls."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ299060; CAC10209.1; -;
 FT NON_TER
 SO SEQUENCE 369 AA; 39776 MW; 5DF6FCLDBF3B212B CRC64;
 Query Match 43.4%; Score 63; DB 10; Length 369;
 Best Local Similarity 50.0%; Pred. No. 1.9;
 Matches 13; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 Oy 2 LASIPRGSTGAGLANGLALPAQVAV 27
 Db 109 LNLGLPRITGILGLARSLSLPTQLA 134
 RESULT 11
 08SIU9 PRELIMINARY: PRT; 402 AA.
 ID 08SIU9
 AC 08SIU9
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative dermal glycoprotein.
 GN P0504E02.10.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 NCBI_TaxID=39947;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC clone:P0504E02."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003269; BAB89709.1; -;
 SO SEQUENCE 402 AA; 41793 MW; E8C84A214D6F6720 CRC64;
 Query Match 43.4%; Score 63; DB 10; Length 402;
 Best Local Similarity 50.0%; Pred. No. 2.1;
 Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
 Oy 1 LLASIPRGSTGAGLANGLALPAQVAS 28
 Db 147 LRLSLPAMAGDAGLGRGVSLPTQLYS 174
 RESULT 12
 09ZVVS PRELIMINARY: PRT; 434 AA.
 ID 09ZVVS
 AC 09ZVVS
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F15K9.16 protein (Putative extracellular dermal glycoprotein precursor).
 GN F15K9.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S., Kremenetskaia I., Lueros J., Araujo R., Buehler E., Conway A.B., Dewar K., Feng J., Kim G., Li Y., Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Egu P., Lee J.M., Toriumi M., Yu G., Brooks S., Chao O., Chen H., Karlin-Neumann G., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F15K9.16 (GI:3850580)."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005278; AAC72120.1; -;
 DR EMBL; AF332411; AAG48774.1; -;
 SO SEQUENCE 434 AA; 46148 MW; 17DD684008FDAFC CRC64;
 Query Match 42.8%; Score 62; DB 10; Length 434;
 Best Local Similarity 41.4%; Pred. No. 3.1;
 Matches 12; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

RC	STRAIN=CV.	NIPPONBARE;
RA	Sasaki T.	Matsumoto T.
		Yamamoto K.

Job time : 40.3242 secs

Job time : 40.3242 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 14.3077 seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-16

Perfect score: 145

Sequence: 1 LASLPRGSTGVAGLANSGIALPQAVASAK 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	37.9	338	4	US-09-325-932A-60
2	48	33.1	208	1	US-08-309-512-4
3	48	33.1	208	1	PCT-US92-08756A-4
4	47.5	32.8	2186	2	US-08-823-445-2
5	47.5	32.8	2186	4	US-09-396-540-2
6	47	32.4	436	2	US-08-958-642-4
7	47	32.4	436	3	US-08-778-394-2
8	47	32.4	436	3	US-08-778-423A-4
9	46.5	32.1	330	3	US-08-115-753-1
10	46.5	32.1	364	4	US-09-333-423-4
11	46.5	32.1	419	3	US-08-115-753-2
12	46.5	32.1	419	3	US-08-115-753-3
13	46	31.7	440	6	5310667-9
14	46	31.7	442	1	US-08-476-008-64
15	46	31.7	442	1	US-08-306-063-64
16	46	31.7	442	1	US-08-833-485-64
17	46	31.7	442	1	US-09-137-440-64
18	45.5	31.4	456	2	US-08-910-731-4
19	45.5	31.4	456	2	US-08-795-395-4
20	45.5	31.4	3672	2	US-08-822-445-12
21	45.5	31.4	3672	4	US-09-396-540-12
22	45.5	31.4	3801	2	US-08-822-445-10
23	45.5	31.4	3801	4	US-09-396-540-10
24	45	31.0	638	4	US-09-070-637-20
25	45	31.0	731	2	US-08-911-364-1
26	45	31.0	733	4	US-08-464-700-2
27	45	31.0	792	2	US-08-678-039A-40

28	45	31.0	2265	2	US-08-149-097D-36	Sequence 36, Appl
29	45	31.0	2509	2	US-08-149-097D-35	Sequence 35, Appl
30	45	31.0	3739	3	US-09-320-878-2	Sequence 2, Appl
31	45	31.0	3739	4	US-09-105-537-33	Sequence 33, Appl
32	45	31.0	11877	4	US-09-105-537-6	Sequence 6, Appl
33	44.5	30.7	220	4	US-09-172-952-28	Sequence 28, Appl
34	44.5	30.7	468	4	US-08-487-596-8	Sequence 8, Appl
35	44.5	30.7	1115	3	US-08-323-477-2	Sequence 2, Appl
36	44	30.3	201	2	US-08-911-364-2	Sequence 2, Appl
37	44	30.3	384	1	US-08-707-793A-5	Sequence 5, Appl
38	44	30.3	384	1	US-08-707-792A-5	Sequence 5, Appl
39	44	30.3	449	4	US-09-230-371A-28	Sequence 28, Appl
40	44	30.3	2584	3	US-08-936-135-4	Sequence 4, Appl
41	43.5	30.0	642	4	US-08-911-393-4	Sequence 4, Appl
42	43.5	30.0	1060	4	US-08-911-393-2	Sequence 2, Appl
43	43.5	30.0	1257	4	US-09-220-641-3	Sequence 3, Appl
44	43.5	30.0	1618	4	US-08-462-467B-4	Sequence 4, Appl
45	43.5	30.0	2588	3	US-08-936-135-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-325-932A-60
; Sequence 60, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE OF INVENTION: death and their use in the modification of forestry plant dev
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325, 932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-60

Query Match      37.9% Score 55; DB 4; Length 338;
Best Local Similarity 44.4%; Pred No. 1.2;
Matches 12; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY      2 LASLPRGSTGVAGLANSGIALPQAVAS 28
      11 1 : |||::: |||:|:|
Db      153 LAVSPAVTDGVLGSLSAQVSLPSQLAS 179

RESULT 2
US-08-309-512-4
; Sequence 4, Application US/08309512
; Patent No. 5759828
; GENERAL INFORMATION:
; APPLICANT: Tal, Ronny
; APPLICANT: Benziman, Moshe
; APPLICANT: Gelfand, David H.
; APPLICANT: Ben-Bassat, Arie
; APPLICANT: Calhoon, Roger D.
; APPLICANT: Wong, Hing C.
; TITLE OF INVENTION: CYCLIC DICUANYLATE METABOLIC ENZYMES
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
```



```

US-08-778-423A-4
; Sequence 4, Application US/08778423A
; Patent No. 6071697
; GENERAL INFORMATION:
;   APPLICANT:
;   APPLICANT:
;   APPLICANT:

```

TITLE OF INVENTION: NOVEL METHOD FOR TESTING THE
TITLE OF INVENTION: DIFFERENTIATION STATUS IN PANCREATIC CELLS OF A MAMMAL
NUMBER OF SEQUENCES: 16
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,423A
FILING DATE: December 31, 1996
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-778-423A-4

Query Match 32.4%; Score 47; DB 3; Length 436;
Best Local Similarity 34.5%; Pred. No. 27;
Matches 10; Conservative 8; Mismatches 9; Indels 2; Gaps 1;

QY 4 SLPRGSGV--AGLANSGLAPQVYASQ 30
DB 397 SQPMGTSGTSTGLTSPGVSVQVPGSE 425

RESULT 9
US-08-115-753-1

Sequence 1, Application US/08115753
Patent No. 6017762

GENERAL INFORMATION:

APPLICANT: JARA, Patrick

APPLICANT: LOISON, Gerard

APPLICANT: RAZANAMPARANY, Voahangy

TITLE OF INVENTION: Cassette for the expression of an

TITLE OF INVENTION: endothiapsin precursor in Cryphonectria Parasitica

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: P.O. BOX 299

STREET: King Street Station, Suite 500, 1800 Diagonal Road

CITY: ALEXANDRIA

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/115,753

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/758872

FILING DATE: 07-SEP-1991

APPLICATION NUMBER: FR 9011230

FILING DATE: 11-SEP-1990

ATTORNEY/AGENT INFORMATION:

NAME: Saxe, Bernhard D

REGISTRATION NUMBER: 28,665

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: endothiapsin
US-08-115-753-1

Query Match 32.1%; Score 46.5; DB 3; Length 330;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 12; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

QY 9 STGVAGLANSGLALPQVYASA 29
DB 211 STSIDGLADTGTTLVLPATVYSA 234

RESULT 10
US-09-333-423-4

Sequence 4, Application US/09333423
Patent No. 6265636

GENERAL INFORMATION:

APPLICANT: Randall, Douglas

APPLICANT: Thelen, Jay

APPLICANT: Miernyk, Jan

APPLICANT: Muszynski, Michael

TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase

TITLE OF INVENTION: Polynucleotides, Polypeptides and Uses thereof

FILE REFERENCE: 0818

CURRENT APPLICATION NUMBER: US/09/333,423

CURRENT FILING DATE: 1999-06-15

EARLIER APPLICATION NUMBER: 60/089,998

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 364

TYPE: PRT

ORGANISM: Zea mays

US-09-333-423-4

Query Match 32.1%; Score 46.5; DB 4; Length 364;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 12; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 6 PRGSGVAGLANSGLALPQVYASA 29
DB 175 PKPEGVIGLINTRLS-PIQVADA 197

RESULT 11
US-08-115-753-2

Sequence 2, Application US/08115753
Patent No. 6017762

GENERAL INFORMATION:

APPLICANT: JARA, Patrick

APPLICANT: LEGOUX, Richard

APPLICANT: LOISON, Gerard

APPLICANT: RAZANAMPARANY, Voahangy

TITLE OF INVENTION: Cassette for the expression of an

TITLE OF INVENTION: endothiapsin precursor in Cryphonectria Parasitica

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: P.O. BOX 299

STREET: King Street Station, Suite 500, 1800 Diagonal Road

CITY: ALEXANDRIA

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/115,753
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/758872
FILING DATE: 07-SEP-1991
APPLICATION NUMBER: FR 9011230
FILING DATE: 11-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, BERNHARD D
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 16781/332
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: preproendolthapsin
FEATURE:
NAME/KEY: Protein
LOCATION: 90..419
US-08-115-753-2

Query Match 32.1%; Score 46.5; DB 3; Length 419;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 12; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

OY 9 STGVAGLANS---LALPAQVASA 29
||:|:|:|:| ||||| ||
Db 300 STSIDGIADGTTLTLPLATVSA 323

RESULT 12
US-08-115-753-33
Sequence 33, Application US/08115753
Patent No. 6017762
GENERAL INFORMATION:
APPLICANT: JARA, Patrick
APPLICANT: LEGOUX, Richard
APPLICANT: LOISON, Gerard
APPLICANT: RAZANAMPARANY, Voahangy
TITLE OF INVENTION: Cassette for the expression of an
TITLE OF INVENTION: endolthapsin precursor in Cryphonectria Parasitica
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.O. BOX 299
STREET: King Street Station, Suite 500, 1800 Diagonal Road
CITY: ALEXANDRIA
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/115,753
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/758872
FILING DATE: 07-SEP-1991
APPLICATION NUMBER: FR 9011230
FILING DATE: 11-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, BERNHARD D
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 16781/332
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-115-753-33

Query Match 32.1%; Score 46.5; DB 3; Length 419;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 12; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

OY 9 STGVAGLANS---LALPAQVASA 29
||:|:|:|:| ||||| ||
Db 300 STSIDGIADGTTLTLPLATVSA 323

RESULT 13
5310667-9
Patent No. 5310667
APPLICANT: EICHHOLTZ, DAVID A.; GASSER, CHARLES S.; KISHORE,
GANESH M.
TITLE OF INVENTION: GLYPHOSATE-TOLERANT 5-ENOLPYRUVYL
-3-PHOSPHOSHIKIMATE SYNTHASES
NUMBER OF SEQUENCES: 37
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/380,963
FILING DATE: 17-JUL-1989
SEQ ID NO: 9;
LENGTH: 440
5310667-9

Query Match 31.7%; Score 46; DB 6; Length 440;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 LLASLRGSGTGAAGLANS 18
|||:| ||||:| |:
Db 33 LLAAALGSGTETGLDLS 50

RESULT 14
US-08-476-008-64
Sequence 64, Application US/08476008
Patent No. 5627061
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-64

Query Match 31.7%; Score 46; DB 1; Length 442;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 LLASLPRGSTGVAGLANS 18
|||:| ||| : || : |
Db 33 LLAALAEGETETGLDLS 50

RESULT 15
US-08-306-063-64
Sequence 64, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glycosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611

FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-64

Query Match 31.7%; Score 46; DB 1; Length 442;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 LLASLPRGSTGVAGLANS 18
|||:| ||| : || : |
Db 33 LLAALAEGETETGLDLS 50

Search completed: May 20, 2003, 16:44:41
Job time : 15.3077 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 ; Search time 15.3297 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-16
145
Sequence: 1 LLASLPGRSTGVAGLANSGLALPAQVASACK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCU_NEM_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCUUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	55	37.9	338	9	US-10-219-220-60
2	55	37.9	556	9	US-10-219-220-259
3	48	33.1	319	9	US-09-738-626-5702
4	47.5	32.8	2186	10	US-09-927-668-2
5	46.5	32.1	364	12	US-10-062-254-363
6	46.5	32.1	447	9	US-09-738-626-4148
7	46.5	32.1	1089	9	US-10-174-590-266
8	46.5	32.1	1089	9	US-10-176-758-266
9	46.5	32.1	1089	9	US-10-175-737-266
10	46.5	32.1	1089	9	US-10-173-706-266
11	46.5	32.1	1089	9	US-10-175-738-266
12	46.5	32.1	1089	9	US-10-175-752-266
13	46.5	32.1	1089	9	US-10-176-482-266
14	46.5	32.1	1089	9	US-10-176-757-266
15	46.5	32.1	1089	9	US-10-176-913-266
16	46.5	32.1	1089	9	US-10-180-552-266
17	46.5	32.1	1089	9	US-10-180-557-266
18	46.5	32.1	1089	9	US-10-173-700-266
19	46.5	32.1	1089	9	US-10-174-572-266

20	46.5	32.1	1089	9	US-10-174-579-266	Sequence 266, App
21	46.5	32.1	1089	9	US-10-174-582-266	Sequence 266, App
22	46.5	32.1	1089	9	US-10-174-588-266	Sequence 266, App
23	46.5	32.1	1089	9	US-10-175-739-266	Sequence 266, App
24	46.5	32.1	1089	9	US-10-175-740-266	Sequence 266, App
25	46.5	32.1	1089	9	US-10-175-743-266	Sequence 266, App
26	46.5	32.1	1089	9	US-10-176-488-266	Sequence 266, App
27	46.5	32.1	1089	9	US-10-176-492-266	Sequence 266, App
28	46.5	32.1	1089	9	US-10-176-747-266	Sequence 266, App
29	46.5	32.1	1089	9	US-10-176-750-266	Sequence 266, App
30	46.5	32.1	1089	9	US-10-176-985-266	Sequence 266, App
31	46.5	32.1	1089	9	US-10-176-987-266	Sequence 266, App
32	46.5	32.1	1089	9	US-10-176-991-266	Sequence 266, App
33	46.5	32.1	1089	9	US-10-176-992-266	Sequence 266, App
34	46.5	32.1	1089	9	US-10-176-993-266	Sequence 266, App
35	46.5	32.1	1089	9	US-10-184-638-266	Sequence 266, App
36	46.5	32.1	1089	9	US-10-173-695-266	Sequence 266, App
37	46.5	32.1	1089	9	US-10-173-697-266	Sequence 266, App
38	46.5	32.1	1089	9	US-10-173-705-266	Sequence 266, App
39	46.5	32.1	1089	9	US-10-174-576-266	Sequence 266, App
40	46.5	32.1	1089	9	US-10-174-585-266	Sequence 266, App
41	46.5	32.1	1089	9	US-10-174-586-266	Sequence 266, App
42	46.5	32.1	1089	9	US-10-175-747-266	Sequence 266, App
43	46.5	32.1	1089	9	US-10-176-481-266	Sequence 266, App
44	46.5	32.1	1089	9	US-10-176-485-266	Sequence 266, App
45	46.5	32.1	1089	9	US-10-176-487-266	Sequence 266, App

ALIGNMENTS

RESULT 1
US-10-219-220-60
; Sequence 60, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. NO. US20030082724A1 09/325,932
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-60

Query Match 37.9%; Score 55; DB 9; Length 338;
Best Local Similarity 44.4%; Pred No. 7;
Matches 12; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 LLASLPGRSTGVAGLANSGLALPAQVAS 28
DB 153 LAVSPAVTGVGLSSAQVSLPSQLAS 179

RESULT 2
US-10-219-220-259
; Sequence 259, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220

CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 290
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 259
LENGTH: 556
TYPE: PRT
ORGANISM: Pinus radiata
US-10-219-220-259

Query Match 37.9%; Score 55; DB 9; Length 556;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 12; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 LASLPGSTGVAGLANGLALPAQVAS 28
DB 309 LAVSPAVTDVGLGSSAQVSLPSQLAS 335

RESULT 3

US-09-738-626-5702
Sequence 5702, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5702
LENGTH: 319
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5702

Query Match 33.1%; Score 48; DB 9; Length 319;
Best Local Similarity 48.3%; Pred. No. 59;
Matches 14; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

QY 2 LASLP--RGSTGVAGLANGLALPAQVAS 28
DB 151 LASLAQVRGVTGPAGAFASDVLPSTVA 179

RESULT 4

US-09-927-668-2
Sequence 2, Application US/09927668
Patent No. US20020115144A1
GENERAL INFORMATION:
APPLICANT: Kaplan, Jerry
Perou, Charles

TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/927,668
FILING DATE: 10-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/396,540
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 2186 amino acids
TYPE: amino acid

MOLECULE TYPE: protein

FRAGMENT TYPE: Internal

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-927-668-2
Query Match 32.8%; Score 47.5; DB 10; Length 2186;
Best Local Similarity 50.0%; Pred. No. 5,4e+02;
Matches 14; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 LASLPGSTGVAGLANGLALPAQVASA 29
DB 621 LASPQRSQSTVASL--GLAFPSQNSGA 645

RESULT 5

US-10-062-254-363
Sequence 363, Application US/10062254
Patent No. US20020138882A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl

APPLICANT: Fang, Yiyen
APPLICANT: Hantke, Sabine S.

APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen

APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele

APPLICANT: Ni, Xiping
APPLICANT: Odell, Joan

APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime

APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/062,254
CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/146511

OY 5 LPRGSTGVAGLANSGLAL 22
||| ||||| |||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 9

US-10-175-737-266
; Sequence 266, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jjian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C50
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-19
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 266
LENGTH: 1089
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-266

Query Match 32.1%; Score 46.5; DB 9; Length 1089;
Best Local Similarity 72.2%; Pred. No. 3.5e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
||| ||||| |||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 10

US-10-173-706-266
; Sequence 266, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jjian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
CURRENT FILING DATE: 2002-06-17
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 266
LENGTH: 1089
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-266

Query Match 32.1%; Score 46.5; DB 9; Length 1089;
Best Local Similarity 72.2%; Pred. No. 3.5e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
||| ||||| |||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 11

US-10-175-738-266
; Sequence 266, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jjian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-19
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 266
LENGTH: 1089
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-266

Query Match 32.1%; Score 46.5; DB 9; Length 1089;
Best Local Similarity 72.2%; Pred. No. 3.5e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
||| ||||| |||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 12

US-10-175-752-266
; Sequence 266, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jjian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C60
CURRENT APPLICATION NUMBER: US/10/175,752
CURRENT FILING DATE: 2002-06-19
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 266
LENGTH: 1089
TYPE: PRT


```
; ORGANISM: Homo Sapien
US-10-175-752-266

Query Match
Best Local Similarity 72.2%; DB 9; Length 1089;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
   ||| |||| ||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 13
US-10-176-482-266
; Sequence 266, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 266
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-266

Query Match
Best Local Similarity 72.2%; DB 9; Length 1089;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
   ||| |||| ||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 14
US-10-176-757-266
; Sequence 266, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
```

```
; SEQ ID NO 266
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-266

Query Match
Best Local Similarity 72.2%; DB 9; Length 1089;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
   ||| |||| ||||
Db 743 LPR---AVAGLAASGLAL 757

RESULT 15
US-10-176-913-266
; Sequence 266, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 266
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-266

Query Match
Best Local Similarity 72.2%; DB 9; Length 1089;
Matches 13; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 5 LPRGSTGVAGLANSGLAL 22
   ||| |||| ||||
Db 743 LPR---AVAGLAASGLAL 757

Search completed: May 20, 2003, 18:00:02
Job time : 16.3297 secs
```

THIS PAGE BLANK (uspro)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 ; Search time 26.1099 Seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-17

Perfect score: 124

Sequence: 1 GGSPAHYISARFIEVGDTRVPSVE 24

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002.*
1: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	24	21	AAV93760
2	124	100.0	24	22	AAE14691
3	124	100.0	24	22	AAU07396
4	112	90.3	381	23	AAU75813
5	105	84.7	381	23	AAU75814
6	104	83.9	381	23	AAU75812
7	101	81.5	196	23	AAU75817
8	97	78.2	370	23	AAU75815
9	97	78.2	370	23	AAU75816
10	75	60.5	185	23	AAU75824

11	75	60.5	185	23	AAU75825	Barley L endoxylan
12	75	60.5	185	23	AAU75826	Barley L endoxylan
13	75	60.5	277	23	AAU75818	L endoxylanase inh
14	65	52.4	102	23	AAU75820	Rye L endoxylanase
15	65	52.4	102	23	AAU75821	Rye L endoxylanase
16	61	49.2	287	23	AAU75819	Wheat L endoxylanase
17	51	41.1	170	23	AAU75827	Wheat L endoxylanase
18	49	39.5	218	22	ABB71510	Oat L endoxylanase
19	48	38.7	410	13	AAU28030	Drosophila melanog
20	48	38.7	862	22	AAU00022	Pichia pastoris pr
21	47	37.9	366	23	ABP40222	Human activated T
22	45	36.3	89	19	AAW42021	Staphylococcus epi
23	45	36.3	89	20	AAU08629	Clone D305_2 prote
24	45	36.3	89	21	AAU67320	Human secreted pro
25	45	36.3	540	22	ABB65612	Drosophila melanog
26	45	36.3	1201	20	AAW90345	Drosophila sp. Cos
27	45	36.3	1201	20	ABB58421	Drosophila melanog
28	45	36.3	1348	22	ABB63571	Drosophila melanog
29	44	35.5	122	23	ABP01617	Human ORFX protein
30	44	35.5	208	21	AAU85644	Oxidoreductase ami
31	44	35.5	329	21	AAU04698	Arabidopsis thalia
32	44	35.5	485	21	AAU04697	Arabidopsis thalia
33	44	35.5	496	21	AAU04696	Arabidopsis thalia
34	43.5	35.1	70	22	AAU28784	Peptide #2821 enco
35	43.5	35.1	70	23	ABG38068	Human peptide enco
36	43	34.7	54	22	AAU66550	Propionibacterium
37	43	34.7	181	23	ABB53437	Lactococcus lactis
38	43	34.7	190	22	ABB27458	Human peptide #109
39	43	34.7	190	22	ABB32608	Peptide #114 enco
40	43	34.7	190	22	ABB18104	Protein #103 enco
41	43	34.7	190	22	AAU53437	Human brain expres
42	43	34.7	190	22	AAU65815	Human bone marrow
43	43	34.7	190	22	AAU13677	Peptide #111 enco
44	43	34.7	190	22	AAU26076	Peptide #113 enco
45	43	34.7	190	22	AAU01427	Peptide #109 enco

ALIGNMENTS

RESULT 1	AAV93760	standard; peptide; 24 AA.
ID	AAV93760	
AC	AAV93760;	
DT	03-OCT-2000	(first entry)
DE	Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.	
XX	Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;	
KW	dough; dough preparation.	
OS	Triticum sp.	
XX		
PN	WO200039289-A2.	
PD	06-JUL-2000.	
XX		
PF	17-DEC-1999;	99WO-1B02071.
XX		
PR	23-DEC-1998;	98GB-0028599.
PR	06-APR-1999;	99GB-0007805.
PR	15-APR-1999;	99GB-0008645.
PA	(DANI-) DANISCO AS.	
XX		
PI	Sibbesen O. Sorensen JF;	
XX		
DR	WPI; 2000-465744/40.	
XX		
PT	Mutant xylanase protein identified using xylanase inhibitor useful for preparing non-sticky dough for bakery products -	

XX Claim 24; Page 112; 112pp; English.
PS
XX The present sequence is derived from an endo-beta-1,4-xylinase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xylinase protein. The xylinase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xylinase or
CC mutant xylinase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xylinase. The xylinase inhibitor is
CC useful for screening high degree resistance xylinases for dough
CC preparation. The xylinase is also useful for preparing a non-sticky
CC dough. A combination of xylinase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
SQ Sequence 24 AA:
Query Match 100.0%; Score 124; DB 21; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.3e-14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSPAHYISARFIEGDRVPSVE 24
Db 1 GSPAHYISARFIEGDRVPSVE 24
RESULT 2
ID AAE14691 standard; peptide: 24 AA.
XX AAE14691:
XX 21-AUG-2002 (first entry)
XX
XX Wheat flour xylinase inhibitor B chain Lys-C digested fragment #2.
DE
XX Refrigerated dough; syrruping; arabinoxylan; bakery product; bread;
KW pizza base; cake; biscuit; wheat; flour; xylinase inhibitor.
XX
XX Triticum aestivum.
OS
XX WO200152657-A1.
PN
XX 26-JUL-2001.
PD
XX 17-JAN-2001; 2001WO-1B00168.
PE
XX 18-JAN-2000; 2000GB-0001136.
PR
XX (DANI-) DANISCO AS.
PA
XX Poulsen CH, Sorensen JF;
PI
XX WPI: 2001-457446/49.
DR
XX
XX Production of refrigerated dough with reduced syrruping, useful in
PT production of bakery products such as bread, comprises admixing cereal
PT flour, water and protein that prevents enzymatic degradation of
PT arabinoxylan in the cereal flour -
XX
XX Disclosure; Page 23; 26pp; English.
PS
XX The invention relates to a process for producing refrigerated dough
CC with reduced 'syrruping' (precipitation of liquid on the dough surface
CC because of a reduction in water holding capacity caused by the breakdown
CC of arabinoxylan over time). The process comprises admixing cereal flour
CC and water with a protein that reduces/prevents enzymatic degradation of
CC arabinoxylan in the cereal flour. The preferred protein is a xylinase
CC inhibitor. The method is useful to produce refrigerated dough in which
CC syrruping is reduced or eliminated. Refrigerated dough is typically
CC stored for long periods to enable fresh baked products (e.g. bread,
CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
CC requirements using the method by the use of specific proteins/protein
CC combinations. The present sequence is wheat flour
CC endo-beta-1,4-xylinase inhibitor B chain Lys-C digested fragment.
XX
SQ Sequence 24 AA:
Query Match 100.0%; Score 124; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.3e-14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSPAHYISARFIEGDRVPSVE 24
Db 1 GSPAHYISARFIEGDRVPSVE 24
RESULT 3
ID AAU07396 standard; protein: 24 AA.
XX AAU07396:
XX 18-DEC-2001 (first entry)
XX
XX Bacillus subtilis xylinase inhibitor #5.
DE
XX Xylinase; plant cell wall; baking; cereal; starch production; wood;
KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
XX
XX Bacillus subtilis.
OS
XX WO20016711-A1.
PN
XX 13-SEP-2001.
PD
XX 08-MAR-2001; 2001WO-1B00426.
PE
XX 08-MAR-2000; 2000GB-0005585.
PR 27-JUN-2000; 2000GB-0015751.
XX
XX (DANI-) DANISCO AS.
PA
XX Sidsesen O, Sorensen JF;
PI
XX WPI: 2001-596834/67.
DR
XX
XX Novel variant xylinase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xylinase inhibitor -
XX
XX Disclosure; Page 63; 70pp; English.
PS
XX The invention relates to a variant xylinase polypeptide (I) or its
CC fragment having xylinase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xylinase inhibitor as compared with the parent xylinase enzyme. (I) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I)
CC is useful for altering the viscosity derived from the presence of
CC hemicellulose or arabinoxylan in a solution or system comprising plant
CC cell wall material. (I) is useful for preparing a foodstuff such as
CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
CC The present sequence represents the amino acid sequence of Bacillus
CC subtilis xylinase inhibitor #5 as described in the method of the
XX invention.
XX
SQ Sequence 24 AA:
Query Match 100.0%; Score 124; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.3e-14;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGSPAHYISARFIEVGDTRPVSE 24
 |||||||||||||||||||
 Db 1 GGSPAHYISARFIEVGDTRPVSE 24

RESULT 4
 AAU75813
 ID AAU75813 standard; Protein: 381 AA.

XX AC AAU75813;
 DT 23-APR-2002 (first entry)
 XX DE Wheat L endoxylanase inhibitor, TAXI I, variant #2.

XX KW Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX OS Triticum aestivum.
 XX FH Key Location/Qualifiers
 FT Misc-difference 95 /label= unknown
 FT Misc-difference 98 /label= unknown
 FT Misc-difference 101 /label= unknown
 FT Misc-difference 110 /label= unknown
 FT Misc-difference 110 /label= unknown
 FT Misc-difference 333 /label= unknown
 FT Misc-difference 333 /label= unknown

XX PN WO200198474-A1.
 XX PD 27-DEC-2001.
 XX PF 21-JUN-2001; 2001WO-BE00106.
 XX PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX PA (LEUV-) LEUVEN RES 6 DEV.
 XX PI Delcours J, Debysse W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Campenhout S;
 DR WPI: 2002-114579/15.

XX PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX PS Claim 127; Page 51; 127pp; English.

XX CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,

CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic acid
 CC acid, a host organism transformed with the nucleic acid of the proteins.
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruiping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-I ('T. aestivum L endoxylanase
 CC inhibitor').
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

XX SQ Sequence 381 AA:
 XX
 XX Query Match 90.3%; Score 112; DB 23; Length 381;
 XX Best Local Similarity 91.7%; Pred. No. 1.7e-10;
 XX Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGSPAHYISARFIEVGDTRPVSE 24
 |||||||||||||||||||
 Db 201 GGSPAHYISARFIEVGDTRPVSE 224

RESULT 5
 AAU75814
 ID AAU75814 standard; Protein: 381 AA.

XX AC AAU75814;
 XX DT 23-APR-2002 (first entry)

XX DE Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.

XX KW Wheat; TAXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX OS Triticum aestivum.
 XX FH Key Location/Qualifiers
 FT Misc-difference 95 /label= unknown
 FT Misc-difference 98 /label= unknown
 FT Misc-difference 101 /label= unknown
 FT Misc-difference 110 /label= unknown
 FT Misc-difference 110 /label= unknown
 FT Misc-difference 145 /label= unknown
 FT Misc-difference 183 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 183 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 232 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 275 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys

QY	1	GGSPAHYISARFIEVGDRVPSVE	24
XX	Sequence	381 AA:	
XX	Query Match	84.7%;	Score 105; DB 23; Length 381;
XX	Best Local Similarity	87.5%;	Pred. No. 2.7e-09;
XX	Matches	21; Conservative	0; Mismatches 3; Indels 0; Gaps 0;

RESULT 6
 ID AAN75812
 XX AAN75812 standard; Protein; 381 AA.
 AC AAN75812;
 XX
 DT 23-APR-2002 (first entry)
 XX
 XX Wheat L endoxylanase inhibitor, TAXI I, variant #1.
 XX
 KW Wheat: TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough sytruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 XX plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 OS Triticum aestivum.
 XX
 FH Key Location/Qualifiers
 FH Misc-difference 95
 FT /label= unknown
 FT Misc-difference 98
 FT /label= unknown
 FT Misc-difference 101
 FT /label= unknown
 FT Misc-difference 110
 FT /label= unknown
 FT Misc-difference 333
 FT /label= unknown
 XX
 XX WO200198474-A1.
 PD
 XX
 PD 27-DEC-2001.
 XX
 PD 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 PR
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI
 PI Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI; 2002-114579/15.
 XX
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 50-51; 127p; English.
 XX
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabinofuranosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-gluacan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory

CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-1 (T. aestivum L. endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

CC Sequence 381 AA:

Query Match 83.9%; Score 104; DB 23; Length 381;
 Best Local Similarity 87.5%; Pred. No. 3.9e-09;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGSPAHYISARFIEVGDTRVPVSE 24
 |||||

Db 201 GGSPAHYISARFIEVGDTRVPVSE 224

RESULT 7

AAU75817
 ID AAU75817 standard; Protein; 196 AA.

AC AAU75817;

DT 23-APR-2002 (first entry)

DE Wheat L endoxylanase inhibitor, TDXI I, partial sequence TDXI-I.01.

KW Wheat; TDXI-I; L endoxylanase inhibitor; cellulosytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Triticum durum cultivar Mexicali.

PN WO200198474-A1.

PD 27-DEC-2001.

PF 21-JUN-2001; 2001WO-BE00106.

PR 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.

PR 16-MAR-2001; 2001GB-0006564.

PR 21-MAY-2001; 2001GB-0012328.

PA (LEUV-) LEUVEN RES & DEV.

PI Delcourt J, Debysier W, Gebuwers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;

DR WPI: 2002-114579/15.

DR N-PSDB; ABK13674.

PT Separating and/or isolating inhibitors of cellulosytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes

PS Claim 127; Page 57; 127pp; English.

CC The invention relates to separating and/or isolating inhibitors of
 CC cellulosytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylidase,
 CC alpha-L-arabino-furanosidase and/or other cellulase, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, a host organism transformed with the nucleic acid, the inhibitory
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, wheat xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TDXI-I (T. durum L. endoxylanase
 CC inhibitor).

SO Sequence 196 AA:

Query Match 81.5%; Score 101; DB 23; Length 196;
 Best Local Similarity 87.0%; Pred. No. 5.7e-09;

Matches 20; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGSPAHYISARFIEVGDTRVPV 23
 |||||

Db 15 GGSPAHYISARFIEVGDTRVPV 37

RESULT 8

AAU75815
 ID AAU75815 standard; Protein; 370 AA.

AC AAU75815;

DT 23-APR-2002 (first entry)

DE Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.

KW Wheat; TAXI-I; L endoxylanase inhibitor; cellulosytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Triticum aestivum cultivar Soissons.

PN WO200198474-A1.

PD 27-DEC-2001.

```

PE 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
PR 25-JAN-2001; 2001GB-0002018.
PR 26-JAN-2001; 2001GB-0002194.
PR 16-MAR-2001; 2001GB-0006564.
PR 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES 6 DEV.
PA
XX Delcours J, Delyser W, Gebruers K, Goesaert H, Flereus K, Robben J;
PI Van Campenhout S;
XX
DR WPI: 2002-114579/15.
DR N-PSDB: ABK13672.
XX
XX Separating and/or isolating inhibitors of cellulolytic, xyanoalytic, or
PT beta-glucanolytic enzymes comprises using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
XX Claim 127; Page 57; 127pp; English.
XX
XX The invention relates to separating and/or isolating inhibitors of
CC cellulolytic, xyanoalytic and/or beta-glucanolytic enzymes comprises
CC screening the inhibition activity by using two or more enzymes during the
CC separation and/or isolation steps that allow to distinguish inhibitors of
CC different specificity or by using an affinity chromatographic step with
CC immobilised enzymes and/or antibodies against inhibitors. Also
CC included are an isolated nucleic acid molecule encoding an inhibitor
CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
CC arabinoxylan or beta-galcan degrading enzymes, a vector comprising the
CC nucleic acid, an expression system transformed with the nucleic
CC acid, a host organism transformed with the nucleic acid, the inhibitory
CC proteins encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syruup in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specificity and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC arabinoxylanase/inhibitor complex. Improving the maling of cereals such as
CC barley, sorgum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
CC starch separation and production, improving maize processing.
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
XX inhibitor).
XX
XX Sequence 370 AA:
XX
XX Query Match 78.2%; Score 97; DB 23; Length 370;
XX Best Local Similarity 83.3%; Pred. No. 66-08;
XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0
XX
XX 1 GGSPAHYISARFTEGDTKPVSE 24
XX ||||||||| ||||||| |
XX
XX 190 GGSPAHYISAKTIVGDTKPVSE 213
XX
XX RESULT 9
XX ID AAU75816
XX AC AAU75816;

```

XX 23-APR-2002 (first entry)
XX
XX Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.02.
XX
XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
XX xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
XX immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
XX straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
XX noodle; animal feed; starch separation; maize processing; malting;
XX plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
XX Trifolium aestivum cultivar Estica.
XX
XX WO2001098474-A1.
XX
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
XX 25-JAN-2001; 2001GB-0002018.
XX 26-JAN-2001; 2001GB-0002194.
XX 16-MAR-2001; 2001GB-0006564.
XX 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES & DEV.
XX
XX Delcour J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
XX Van Campenhout S;
XX
XX WPI: 2002-114579/15.
XX N-PSDB; ABRK3673.
XX
XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
XX beta-glucanolytic enzymes comprises using endoxylanases during
XX screening for inhibition activity or affinity chromatography with
XX immobilised enzymes -
XX
XX Claim 127: Page 57; 127pp: English.
XX
XX The invention relates to separating and/or isolating inhibitors of
XX cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
XX screening the inhibition activity by using two or more enzymes during the
XX separation and/or isolation steps that allow to distinguish inhibitors of
XX different specificity or by using an affinity chromatographic step with
XX immobilised enzymes and/or antibodies against inhibitors. Also
XX included are an isolated nucleic acid molecule encoding an inhibitor
XX which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
XX alpha-L-arabinofuranosidase and/or other cellulase, xylan,
XX arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
XX nucleic acid, an expression system transformed with the nucleic
XX acid, a host organism transformed with the nucleic acid, the inhibitory
XX proteins encoded by the nucleic acids and modulators of the proteins.
XX A recombinant protein, glycoprotein or polypeptide or microorganisms,
XX plant or plant materials transformed with the nucleic acid are useful
XX for the formation of an endoxylanase-inhibitor complex, screening
XX endoxylanases that are totally, less or not inhibited by the inhibitors,
XX reducing syripping in refrigerated dough compositions, affecting the
XX relative affinity and/or relative hydrolysis specificity and/or relative
XX hydrolysis rate versus water-extractable and/or water-unextractable
XX arabinoxylans of endoxylanases such as by the formation of an
XX endoxylanase/inhibitor complex, improving the malting of cereals such as
XX barley, sorghum and wheat and/or the production of beer, improving the
XX production and/or quality of baked or extruded cereal products such as
XX straight dough, sponge dough, Chorleywood bread, breakfast cereals,
XX biscuits, pasta and noodles, animal feed stuff, improving the production
XX of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
XX starch separation and production, improving maize processing,
XX plant disease resistance and nutraceutical and/or pharmaceutical
XX applications, improving paper and pulp technologies. The present
XX sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
XX inhibitor).

XX Sequence 370 AA;
SQ Query Match 78.2%; Score 97; DB 23; Length 370;
Best Local Similarity 83.3%; Pred. NO. 6e-08;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 GSPAHYISARFIEVGDPTRVPE 24
||||||| |
Db 190 GSPAHYISARFIEVGDPTRVPE 213

RESULT 10
AAU75824
ID AAU75824 standard; Protein: 185 AA.
XX AAU75824;
AC
XX 23-APR-2002 (first entry)
DT
XX
XX Barley L endoxylanase inhibitor, HVXI I, variant #1.
DE
XX
KW Barley; HVXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
XX Hordeum vulgare.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 72 /label= Unknown
FT Misc-difference 185 /label= Unknown
FT
XX
XX WO200198474-A1.
PN
XX
XX 27-DEC-2001.
PD
XX
XX 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
PR 25-JAN-2001; 2001GB-0002018.
PR 26-JAN-2001; 2001GB-0002194.
PR 16-MAR-2001; 2001GB-0006564.
PR 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES & DEV.
PA
XX
XX Delcours J, Debysse W, Gebuereers K, Goesaert H, Fierens K, Robben J;
PI Van Campenhout S;
XX
XX WPI: 2002-114579/15.
DR
XX
XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT beta-glucanolytic enzymes comprising using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
XX Claim 127; Page 63; 127pp; English.
PS
XX
XX The invention relates to separating and/or isolating inhibitors of
CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
CC screening the inhibition activity by using two or more enzymes during the
CC separation and/or isolation steps that allow to distinguish inhibitors of
CC different specificity or by using an affinity chromatographic step with
CC immobilised enzymes and/or antibodies against inhibitors. Also
CC included are an isolated nucleic acid molecule encoding an inhibitor
CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the

CC nucleic acid, an expression system transformed with the nucleic
CC acid, a host organism transformed with the nucleic acid, the inhibitory
CC proteins encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syruiping in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specificity and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC endoxylanase/inhibitor complex, improving the malting of cereals such as
CC barley, sorghum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylitol, maltitol, wheat gluten
CC starch separation and production, improving maize processing,
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence is a variant of barley HVXI-I (H. vulgare L endoxylanase
CC inhibitor).
CC Note: Variant amino acids are highlighted in the specification but
CC no wild-type sequence is shown for comparison.
XX
SQ Sequence 185 AA;
Query Match 60.5%; Score 75; DB 23; Length 185;
Best Local Similarity 70.0%; Pred. NO. 0.00015;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 2 GSPAHYISARFIEVGDPTRVP 21
||||| | | | |
Db 62 GSPAHYISARFIEVGDPTRVP 81

RESULT 11
AAU75825
ID AAU75825 standard; Protein: 185 AA.
XX
XX AAU75825;
AC
XX
XX 23-APR-2002 (first entry)
DT
XX
XX Barley L endoxylanase inhibitor, HVXI I, variant #2.
DE
XX
KW Barley; HVXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KW immobilised enzyme; enzyme; dough syruiping; cereal product; beer;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
XX Hordeum vulgare.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 72 /label= Unknown
FT Misc-difference 185 /label= Unknown
FT
XX
XX WO200198474-A1.
PN
XX
XX 27-DEC-2001.
PD
XX
XX 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
PR 25-JAN-2001; 2001GB-0002018.
PR 26-JAN-2001; 2001GB-0002194.
PR 16-MAR-2001; 2001GB-0006564.
PR 21-MAY-2001; 2001GB-0012328.
XX

Search completed: May 20, 2003, 16:14:42
Job time : 27.109 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 9.75824 Seconds

(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-17

Perfect score: 124

Sequence: 1 GGSPAHYISARFIEVGDTRVPSVE 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries
PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	48	38.7	136	2 AH3394	lactoylgutathione
2	48	38.7	357	2 F82878	XAA-PRO aminopepti
3	48	38.7	682	2 F70421	conserved hypotet
4	47	37.9	161	2 AE0357	conserved hypotet
5	46	37.1	344	2 A95402	probable desaturas
6	46	37.1	444	2 T38760	hypothetical prote
7	46	37.1	455	2 C83494	probable 2-isoprop
8	46	37.1	486	2 A82427	pyruvate kinase II
9	46	37.1	518	2 A13534	probable binding p
10	46	37.1	634	1 S24384	nitrous-oxide redu
11	46	37.1	682	2 T39613	hypothetical prote
12	46	37.1	714	2 T35770	hypothetical prote
13	45.5	36.7	410	2 C84205	hypothetical prote
14	45	36.3	231	2 F70471	hypothetical prote
15	45	36.3	389	2 AD1918	conserved hypotet
16	45	36.3	625	2 B16555	alcohol dehydrogen
17	45	36.3	852	2 T28790	succinate dehydrog
18	45	36.3	955	2 E84022	hypothetical prote
19	45	36.3	1201	2 T08603	hypothetical prote
20	44	35.5	202	2 T46586	kinesin-related pr
21	44	35.5	314	2 T2383	ribosomal protein
22	44	35.5	414	2 E6387	hypothetical prote
23	44	35.5	485	2 C86143	hypothetical prote
24	44	35.5	592	2 T07616	probable beta-fruc
25	44	35.5	636	2 C83222	nitrous-oxide redu
26	44	35.5	918	2 S44769	hypothetical prote
27	43.5	35.1	4930	2 E86679	C25B4.2 protein -
28	43	34.7	148	2 H90257	polypeptide synthet
29	43	34.7	165	2 A41072	hypothetical prote
					photosystem I chai

30	43	34.7	181	2 F86641	hypothetical prote
31	43	34.7	247	2 G87472	hypothetical prote
32	43	34.7	282	1 E69906	conserved hypotet
33	43	34.7	459	1 DEPSLP	dihydrolipeamide d
34	43	34.7	464	2 F83365	lipamide dehydrog
35	43	34.7	471	2 S30585	hypothetical prote
36	43	34.7	527	2 E70543	probable L-asparta
37	43	34.7	746	2 T24978	hypothetical prote
38	43	34.7	817	2 T03852	hypothetical prote
39	42.5	34.3	412	1 G70017	protein phosphatas
40	42.5	34.3	781	2 T35029	probable N-carbamy
41	42	33.9	119	2 F42523	hypothetical prote
42	42	33.9	224	2 T10120	A-ORF-A protein -
43	42	33.9	356	2 B84183	F420-dependent NAD
44	42	33.9	376	2 D83099	serine proteinase
45	42	33.9	419	2 AE0937	probable RND efflu
					probable serine hy

ALIGNMENTS

```

RESULT 1
AH3394
lactoylgutathione lyase (EC 4.4.1.5) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 22-Mar-2002
C:Accession: AH3394
R:DeVeechlo, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AB3252; PMID:11756688
A:Accession: AH3394
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <KUR>
A:Cross-references: AE008917; PIDN:AAL52323.1; PID:q17963117; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11142
A:Map position: 1
A:Superfamily: hypothetical protein AF2218
C:Keywords: carbon-sulfur lyase

Query Match      38.7%  Score 48:  DB 2:  length 136:
Best Local Similarity 38.1%  Pred. No. 2.6;
Matches      8;  Conservative      7;  Mismatches      6;  Indels      0;  Gaps      0;

OY  4  PAHYISARFIEVGDTRVPSVE 24
Db   39  PERGVTVVFTIDVGNKIELLE 59

RESULT 2
F82878
XAA-PRO aminopeptidase UN532 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: F82878
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views o
A:Reference number: A82870
A:Accession: F82878
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <GLA>
A:Cross-references: GB:AE002152; GB:AF222894; NID:g6899532; PIDN:AAF0945.1; GSP
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: pepP; UN532
A:Genetic code: SGC3
C:Superfamily: X-Pro aminopeptidase

```


A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic path
A:Reference number: AB2950; MWID:20437337; PMID:10984043
A:Accession: C83494
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <STO>
A:Cross-references: GB:AE004551; GB:AE004091; NID:g9947135; PIDN:AAG04606.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1217

Query Match 37.1%; Score 46; DB 2; Length 455;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 GSPAHYISARFIEVGD 17
DB 355 GSPFYIARQFHDVGD 370

RESULT 8

AB2427
pyruvate kinase II VCA0708 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: AB2427
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;
Chadson, D.; Ermolova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AB2035; MWID:20406833; PMID:10952301
A:Accession: AB2427
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-486 <HEI>
A:Cross-references: GB:AE004400; GB:AE003853; NID:g9658121; PIDN:AAF96607.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0708
A:Map position: 2
C:Superfamily: pyruvate kinase

Query Match 37.1%; Score 46; DB 2; Length 486;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 13 IEVGDTRVPSVE 24
DB 253 VEIGDARLPVSQ 264

RESULT 9

AI3534
probable binding protein yddS precursor [imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3534
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
J.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AI3534
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53444.1; PID:g17984342; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10203
A:Map position: 11

Query Match 37.1%; Score 46; DB 2; Length 518;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PAHYISARFIEVGDTR 19
DB 503 PAEYINIRFWEICQAK 518

RESULT 10

S24384
nitrous-oxide reductase (EC 1.7.99.6) [similarity] - *Pseudomonas stutzeri*
C:Species: *Pseudomonas stutzeri*
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: S24384
R:Zunft, W.G.; Dreusch, A.; Lechelt, S.; Cuypers, H.; Friedrich, B.; Schneider, B.
Eur. J. Biochem. 208, 31-40, 1992
A:Title: Derived amino acid sequences of the nosZ gene (respiratory N(2)O reductase)
es. Implications for the Cu(A) site of N(2)O reductase and cytochrome-c oxidase.
A:Reference number: S24382; MWID:92380183; PMID:1324835
A:Accession: S24384
A:Molecule type: DNA
A:Residues: 1-634 <2UM>
A:Cross-references: EMBL:X65277; NID:g45853; PIDN:CAA46381.1; PID:g45855
C:Superfamily: nitrous-oxide reductase
F:579,614,618/Binding site: copper 1 (His, Cys, Cys) #status predicted
F:614,616,618/Binding site: copper 2 (Cys, Trp, Cys, His) #status predicted

Query Match 37.1%; Score 46; DB 1; Length 634;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 SARFIEVGDTRVPSVE 24
DB 285 AGRFITTGDSKVPVVD 300

RESULT 11

T39613
hypothetical protein SPBC16H5.12c - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T39613
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
Submitted to the EMBL Data Library, March 1998
A:Reference number: Z21843
A:Accession: T39613
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-682 <MOO>
A:Cross-references: EMBL:AL022104; PIDN:CAA17910.2; GSPDB:GN00067; SPDB:SPBC16H5.12c
A:Experimental source: strain 972h; cosmid c16H5
C:Genetics:
A:Gene: SPDB:SPBC16H5.12c
A:Map position: 2
A:Introns: 4/1; 41/2
C:Superfamily: *Schizosaccharomyces pombe* hypothetical protein SPBC16H5.12c

Query Match 37.1%; Score 46; DB 2; Length 682;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 12; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 3 SPAHYIS-----ARFIEVGDTRVPSVE 24
DB 248 SPEHYIQLSARAQFMEVYDTVRAEVE 275

RESULT 12

T35770
hypothetical protein SC8A6.03c SC8A6.03c - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C:Accession: T35770
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, July 1998
 A:Reference number: 221570
 A:Accession: T35770
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-714 <SSE>
 A:Cross-references: EMBL:AL031013; PIDN:CAA19775.1; GSPDB:GN00070; SCOPDB:SC8A6.03c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOPDB:SC8A6.03c

Query Match 37.1%; Score 46; DB 2; Length 714;
 Best Local Similarity 42.1%; Pred. No. 35;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 PAHYISARFIEVGDTTPSV 22
 DB 327 PARFVAALAEVGRKRIPN 345

RESULT 13

C84205
 hypothetical protein Vng0468c [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: C84205
 R:N9, W.V.; Kennedy, S.P.; Mahairs, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.; Hough, D.W.; Maddocks, D.G.; Jabido, Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: C84205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-410 <STO>
 A:Cross-references: GB:AE004437; NID:g10580073; PIDN:AA019007.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0468C

Query Match 36.7%; Score 45.5; DB 2; Length 410;
 Best Local Similarity 42.9%; Pred. No. 23;
 Matches 9; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 4 PAHYISARFIEVGDTTPSV 24
 DB 273 PANFTDNFMAIGDT-VPTVD 292

RESULT 14

F70471
 conserved hypothetical protein aq_1997 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000
 C:Accession: F70471
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: F70471
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-231 <NOF>
 A:Cross-references: GB:AE000767; NID:g2984235; PIDN:AA007773.1; PID:g2984245; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: aq_1997
 C:Superfamily: carboxymethylglutaminase

Query Match 36.3%; Score 45; DB 2; Length 231;
 Best Local Similarity 42.9%; Pred. No. 15;
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 GSPAHYISARFIEVGDTTP 21
 DB 120 GGTLMYFAAKPEMVDASLP 140

RESULT 15

AD1918
 alcohol dehydrogenase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AD1918
 R:Keneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriku
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD1918
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-389 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA072852.1; PID:g17130241; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr0895
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match 36.3%; Score 45; DB 2; Length 389;
 Best Local Similarity 45.0%; Pred. No. 26;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GSPAHYISARFIEVGDTTP 21
 DB 137 GAOAEYIRVPFADVGWVKVP 156

Search completed: May 20, 2003, 16:37:04
 Job time : 11.7582 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 6.46154 Seconds
(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-17
Perfect score: 124
Sequence: 1 GGSAPHYISARFIEVGDTRVPSYE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	37.1	444	YAS2_SCHPO	Q10138 schizosacch
2	46	37.1	634	NS22_PSEAE	Q01710 pseudomonas
3	45	36.3	231	DLHM_AQDAE	Q67802 aquifex aeo
4	44	35.5	202	RS7_NEUCR	Q43105 neurospora
5	44	35.5	414	Y701_MERJA	Q58112 methanococ
6	44	35.5	636	NS21_PSEAE	Q09412 pseudomonas
7	44	35.5	821	YK62_CABEL	P34341 caenorhabdi
8	44	35.5	1385	NAC2_CHURE	Q31686 chlamydomon
9	43	34.7	165	PSAF_SYNY3	P39256 synecocyst
10	43	34.7	181	YBDI_LACIA	Q9c672 lactococcus
11	43	34.7	434	AS15_HUMAN	Q8wxk1 homo sapien
12	43	34.7	459	DLDI_PSEPU	P09060 pseudomonas
13	43	34.7	527	NADB_MYCTU	O06595 pseudomonas
14	43	34.7	817	NEB2_RAT	O35274 rattus norv
15	42.5	34.3	412	ALIC_BACCU	P30510 bacillus su
16	42	33.9	119	YVAA_VACCC	O82229 salmonella
17	42	33.9	419	GLA2_SALTI	P58021 mus musculu
18	42	33.9	662	T9S2_MOUSE	Q99805 homo sapien
19	42	33.9	663	T9S2_MOUSE	Q99805 homo sapien
20	42	33.9	663	T9S2_MOUSE	Q99805 homo sapien
21	42	33.9	1135	RBL2_MOUSE	O64700 mus musculu
22	41.5	33.5	436	HEMI_HAUNI	O50581 rattus norv
23	41.5	33.5	436	HEMI_HAUNI	O50581 rattus norv
24	41.5	33.5	436	HEMI_HAUNI	O50581 rattus norv
25	41	33.1	250	LINC_PSEPA	P30668 synecocyst
26	41	33.1	289	AMIA_SALTY	Q9KCY8 bacillus ha
27	41	33.1	299	PTB_BACSU	P30197 pseudomonas
28	41	33.1	390	LPXB_HAEIN	P34530 bacillus su
29	41	33.1	395	KIME_MOUSE	P45011 haemophilus
30	41	33.1	395	PORA_PYRAB	O9U008 mus musculu
31	41	33.1	431	GLAI_RHIME	Q92406 pyrococcus
32	41	33.1	448	ACCC_HAEIN	Q32406 rhizobium m
33	41	33.1	468	GLNA_ECOLI	P43873 haemophilus

34	41	33.1	468	1	GLNA_SALTY	P06201 salmonella
35	41	33.1	505	1	GUAA_PYRAE	O82192 pyrobaculum
36	41	33.1	566	1	ROCB_BACSU	P39635 bacillus su
37	41	33.1	1295	1	BXA2_CLOBO	O45894 clostridium
38	40.5	32.7	331	1	GSPK_AERHY	P31760 aeromonas h
39	40.5	32.7	388	1	GMCR_MOUSE	Q00941 mus musculu
40	40	32.3	125	1	YM07_PARTE	P15608 parametium
41	40	32.3	196	1	RL6_ARCFU	O28370 archaeoglob
42	40	32.3	208	1	FADD_HUMAN	O13158 homo sapien
43	40	32.3	251	1	YN31_PYRAE	O82460 pyrobaculum
44	40	32.3	257	1	NUDC_ECO57	Q8X67 escherichia
45	40	32.3	257	1	NUDC_ECOLI	P32664 escherichia

ALIGNMENTS

RESULT 1
YAS2_SCHPO
ID YAS2_SCHPO STANDARD; PRT; 444 AA.
AC 01-OCR-1996 (Rel. 34, Created)
DT 01-OCR-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C3H8.02 in chromosome I.
GN SPAC3H8.02.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Molyne P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoult B.,
RA Weltjens I., Voiclaert G., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaune V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daea R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesely D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
-1- SIMILARITY: CONTAINS 1 CRAL-TRIO DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: Z69086; CAAG3159.1; -
DR InterPro: IPR001251; CRAL_TRIO.


```

ID RS7_NEUCR STANDARD: PRT: 202 AA.
AC 043105;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 40S Ribosomal Protein S7.
GN CRPS-7.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-IVA;
RX MEDLINE=97174111; PubMed=9021131;
RA Vierula P.J.;
RT "Cloning and characterization of a Neurospora crassa ribosomal
protein gene, crps-7."
RL Curr. Genet. 31:139-143(1997).
CC -1 SIMILARITY: BELONGS TO THE S7E FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: U7847; AAB94301.1;
DR InterPro: IPR000554; Ribosomal_S7E.
DR Pfam: PF01251; Ribosomal_S7e; 1.
DR ProDom: PD006276; Ribosomal_S7E; 1.
DR PROSITE: PS00948; RIBOSOMAL_S7E; 1.
KM Ribosomal protein.
SQ SEQUENCE 202 AA; 22819 MW; 94D87756FE37798 CRC64;

Query Match 35.5%; Score 44; DB 1; Length 202;
Best Local Similarity 37.9%; Pred. No. 6.8;
Matches 11; Conservative 5; Mismatches 5; Indels 8; Gaps 1;

QY 4 PAHYISARFLENGDTR-----VPSVE 24
Db 46 PLOFVSAREIEVGHGKKAIVFVPSLQ 74
1 :||| ||| :
1 :||| ||| :

RESULT 5
Y701_METUA STANDARD: PRT: 414 AA.
AC 058112;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0701.
GN M0701.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayman R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weissflog K.G., Merrick J.M., Nguyen A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.L., Glick A.,
RA Uitterlinden T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";

```

```

RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: U67516; AAB98698.1;
DR TIGR: M0701;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 414 AA; 49101 MW; B53D6FEEB506CAF1 CRC64;

Query Match 35.5%; Score 44; DB 1; Length 414;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGSPAHYISARF1 13
Db 262 GGSPAHYISARFV 274
||| ||| :
||| ||| :

RESULT 6
NSZ1_PSEAE STANDARD: PRT: 636 AA.
AC 09HY12;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nitrous-oxide reductase precursor (EC 1.7.99.6) (N(2)OR) (N2O
reductase).
GN NOSZ OR PA3392.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01.
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: NITROUS-OXIDE REDUCTASE IS PART OF A BACTERIAL
RESPIRATORY SYSTEM WHICH IS ACTIVATED UNDER ANAEROBIC
CONDITIONS IN THE PRESENCE OF NITRATE OR NITROUS OXIDE.
CC -1- CATALYTIC ACTIVITY: N(2) + H(2)O + acceptor = nitrous oxide +
reduced acceptor.
CC -1- COFACTOR: COPPER (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER; EACH SUBUNIT CONTAINS 2 DIINTEGRAL COPPER
CENTERS A AND Z. Z IS THOUGHT TO BE THE SITE OF NITROUS OXIDE
REDUCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic (by similarity).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, TO MITOCHONDRIAL AND
BACTERIAL COX2 SUBUNTS.
CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```


FT PROTEIN ACCUMULATION DECREASES, AND THE
 FT PROTEIN FORMS HIGH MOLECULAR WEIGHT
 FT AGGREGATES.
 SQ SEQUENCE 1385 AA; 144059 MW; 4DB548A508405FE CRC64;
 Query Match 35.5%; Score 44; DB 1; Length 1385;
 Best Local Similarity 42.1%; Pred. No. 59;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 OY 1 GSPAHYISARFIEVGDPTR 19
 Db 950 GGAVMHWGSRLEAGNVR 968
 RESULT 9
 PSFA_SYNV3 STANDARD; PRT: 165 AA.
 AC P29236;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Photosystem I reaction centre subunit III precursor (PSI-F).
 GN PSFA OR SL0819.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92041839; PubMed=1939076;
 RA Chltnis P.R., Purvis D., Nelson N.;
 RT "Molecular cloning and targeted mutagenesis of the gene psaf encoding
 RT subunit III of photosystem I from the cyanobacterium Synechocystis
 RT sp. PCC 6803.";
 RL J. Biol. Chem. 266:20146-20151(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94148808; PubMed=8106355;
 RA Xu Q., Yu L., Chltnis V.P., Chltnis P.R.;
 RT "Function and organization of photosystem I in a cyanobacterial
 RT mutant strain that lacks psaf and psad subunits.";
 RL J. Biol. Chem. 269:3205-3211(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-116(1996).
 CC -I- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER
 CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C55 IN ALGAE AND
 CC CYANOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES
 CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.
 CC -I- SIMILARITY: BELONGS TO THE PSFA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M74801; AAA27292.1; -
 CC EMBL: U20938; AAA27294.1; -
 CC EMBL: D90911; BAA18108.1; -
 CC PIR: A41072; A41072.
 CC InterPro: IPR003666; PSI_Psaf.
 CC Pfam: PF02507; PSI_Psaf.1.

KW Photosynthesis; Photosystem I; Signal; Complete proteome.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 165 PHOTOSYSTEM I REACTION CENTRE SUBUNIT
 FT III.
 SQ SEQUENCE 165 AA; 18249 MW; 9DD84F6B87E72151 CRC64;
 Query Match 34.7%; Score 43; DB 1; Length 165;
 Best Local Similarity 36.4%; Pred. No. 7.9;
 Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
 OY 2 GSPAHYISARFIEVGDPTRPV 23
 Db 71 GXPFLIVDGRFTNAGDFLIPSI 92
 RESULT 10
 YBDI_LACLA STANDARD; PRT: 181 AA.
 ID YBDI_LACLA
 AC Q9CJ67;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ybdi.
 GN YBDI OR L10134.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jellion O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -I- SIMILARITY: BELONGS TO THE UPF0177 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE006251; AAK04232.1; -
 CC InterPro: IPR003675; Abi.
 CC Pfam: PF02517; Abi.1.
 CC KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 SQ SEQUENCE 181 AA; 21601 MW; 7E021BCEC7FCA1 CRC64;
 Query Match 34.7%; Score 43; DB 1; Length 181;
 Best Local Similarity 47.4%; Pred. No. 8.7;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 OY 2 GSPAHYISARFIEVGDPTRV 20
 Db 55 GNTAHYLYRFPVADPTIV 73
 RESULT 11
 AS15_HUMAN STANDARD; PRT: 434 AA.
 ID AS15_HUMAN
 AC O8WKK1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ankyrin repeat and SOCS box containing protein 15 (ASB-15) (Fragment).
 GN ASB15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kille B.T., Nicola N.A.;
 RT 'SOCS box proteins'.
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS AT LEAST 9 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
 CC -----
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR 3MBL: AF403033; AAL57352.1; -
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR001496; SOCS.
 DR Pfam: PF00023; ank; 8.
 DR PRINTS: PR01415; ANKYRIN.
 DR SMART: SM00248; ANK; 7.
 DR SMART: SM00253; SOCS; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 5.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS50225; SOCS; 1.
 DR KW ANK repeat; Repeat.
 FT NON_TER 1 1
 FT REPEAT <1 18 ANK 1.
 FT REPEAT 22 51 ANK 2.
 FT REPEAT 55 84 ANK 3.
 FT REPEAT 88 117 ANK 4.
 FT REPEAT 121 150 ANK 5.
 FT REPEAT 153 182 ANK 6.
 FT REPEAT 195 224 ANK 7.
 FT REPEAT 225 254 ANK 8.
 FT REPEAT 262 290 ANK 9.
 FT DOMAIN 370 425 SOCS BOX.
 FT SO 434 AA; 48347 MW; E69176A4E01D75C6 CRC64;
 SQ SEQUENCE

Query Match 34.7%; Score 43; DB 1; Length 434;
 Best Local Similarity 50.0%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 8 ISARFIEVDTRVPSV 23
 DB 252 VNCYFMHVNDRPPSV 267

RESULT 12
 DDL1_PSEPU
 ID DDL1_PSEPU STANDARD; PRT: 459 AA.
 AC P09063;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Dihydrolipoamide dehydrogenase (EC 1.8.1.4) (E3 component of branched-
 DE chain alpha-keto acid dehydrogenase complex) (LFD-VAL).
 GN LFDV.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PPG2;

RX MEDLINE=88329056; PubMed=3046941;
 RA Burns G., Brown T., Hatter K., Sokatch J.R.;
 RT "Comparison of the amino acid sequences of the transacylase
 RT components of branched chain oxoacid dehydrogenase of Pseudomonas
 RT putida, and the pyruvate and 2-oxoglutarate dehydrogenases of
 RT Escherichia coli.";
 RL Eur. J. Biochem. 176:165-169(1988).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=PPG2.
 RX MEDLINE=89137095; PubMed=2917566;
 RA Burns G., Brown T., Hatter K., Sokatch J.R.;
 RT "Sequence analysis of the lpdv gene for lipamide dehydrogenase of
 RT branched-chain-oxoacid dehydrogenase of Pseudomonas putida.";
 RL Eur. J. Biochem. 179:61-69(1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.45 ANGSTROMS).
 RX MEDLINE=92390345; PubMed=1325638;
 RA Matevi A., Obmolova G., Sokatch J.R., Betzel C., Hol W.G.J.;
 RT "The refined crystal structure of Pseudomonas putida lipamide
 RT dehydrogenase complexed with NAD+ at 2.45-A resolution.";
 RL proteins 13:336-351(1992).
 CC -1- FUNCTION: THE BRANCHED-CHAIN ALPHA-KETO DEHYDROGENASE COMPLEX
 CC CATALYZES THE OVERALL CONVERSION OF ALPHA-KETO ACIDS TO ACYL-COA
 CC AND CO(2). IT CONTAINS MULTIPLE COPIES OF 3 ENZYMIC COMPONENTS:
 CC BRANCHED-CHAIN ALPHA-KETO ACID DECARBOXYLASE (E1), LIPAMIDE
 CC ACYLTRANSFERASE (E2) AND LIPAMIDE DEHYDROGENASE (E3).
 CC -1- CATALYTIC ACTIVITY: Dihydrolipoamide + NAD(+) = lipamide + NADH.
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
 CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
 CC OXIDOREDUCTASES CLASS-1.
 CC -----
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: M57613; AAA65618.1; -
 DR PIR: S02139; DEPSLP.
 DR PDB: 1JVL; 31-JAN-94.
 DR InterPro: IPR001327; FAD_pyr_redox.
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR001100; pyr_redox.
 DR InterPro: IPR004099; pyr_redox_dim.
 DR Pfam: PF00070; pyr_redox; 1.
 DR Pfam: PF02852; pyr_redox_dim; 1.
 DR PRINTS: PR00368; FADPNR.
 DR PRINTS: PR00411; PNDRDTASE1.
 DR ProDom: PD000139; FAD_pyr_redox; 1.
 DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
 KW Redox-active center; Glycolysis; Oxidoreductase; NAD; Flavoprotein;
 KW FAD: 3D-structure.
 FT NP_BIND 8 38 FAD (ADP PART) (PROBABLE).
 FT DISULFID 44 49 REDOX-ACTIVE.
 FT DOMAIN 146 269 NAD-BINDING (BY SIMILARITY).
 FT NP_BIND 296 306 FAD (FLAVIN PART) (BY SIMILARITY).
 FT ACT_SITE 438 438 BY SIMILARITY.
 FT STRAND 5 6
 FT STRAND 9 12
 FT STRAND 16 18
 FT HELIX 16 28
 FT TURN 29 29
 FT STRAND 32 35
 FT TURN 40 41
 FT HELIX 42 47
 FT TURN 49 67
 FT HELIX 68 69
 FT TURN 73 74

DR PRINTS; PRO0411; PNDRTASEI.
 DR TIGR00551; nadb; 1.
 KM Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
 KM Complete proteome.
 SO SEQUENCE 527 AA: 53784 MW: 5494D35AB107ESDA CRC64;

Query Match 34.7%; Score 43; DB 1; Length 527;
 Best Local Similarity 45.5%; Pred. No. 29;
 Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

2 GSPAHYISARFIEGDRPSV 23
 Db 296 GDCPVLDARGIEGFASRPTV 317

RESULT 14
 NEB2_RAT STANDARD; PRT; 817 AA.
 AC 035274;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurabin-II (Neural tissue-specific F-actin binding protein II)
 DE (Protein phosphatase 1 regulatory subunit 9B) (Spinophilin) (P130)
 DE (PPIBP134).
 GN PPIBP9B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97420791; Pubmed-9275233;
 RT Allen P.B., Ouimet C.C., Greengard P.;
 RT "Spinophilin, a novel protein phosphatase 1 binding protein localized
 RT to dendritic spines.";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:9956-9961(1997).
 RL [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-24; 164-182; 199-242; 449-480;
 RP 678-689 AND 776-791.
 RC TISSUE-Brain;
 RX MEDLINE-98123121; Pubmed-9452470;
 RA Satoh A., Nakanishi H., Obaishi H., Wada M., Takahashi K., Satoh K.,
 RA Hiroo K., Nishiohara H., Hata Y., Mizoguchi A., Takai Y.;
 RT "Neurabin-II/spinophilin. An actin filament-binding protein with one
 RT pol domain localized at cadherin-based cell-cell adhesion sites.";
 RL J. Biol. Chem. 273:3470-3475(1998).
 RN [3]
 RP SEQUENCE OF 100-767 FROM N.A., AND INTERACTION WITH D2DR.
 RA MEDLINE-99321921; Pubmed-10391935;
 RA Smith F.D., Oxford G.S., Milgram S.L.;
 RT "Association of the D2 dopamine receptor third cytoplasmic loop with
 RT spinophilin, a protein phosphatase-1-interacting protein.";
 RL J. Biol. Chem. 274:19894-19900(1999).
 RN [4]
 RP CHARACTERIZATION.
 RC TISSUE-Brain;
 RX MEDLINE-20054471; Pubmed-10585469;
 RA Macmillan L.B., Bass M.A., Cheng N., Howard E.F., Tamura M.,
 RA Strack S., Wadzinski B.E., Colbran R.J.;
 RT "Brain actin-associated protein phosphatase 1 holoenzymes containing
 RT spinophilin, neurabin, and selected catalytic subunit isoforms.";
 RL J. Biol. Chem. 274:35845-35854(1999).
 RN [5]
 RP INTERACTION WITH TGN38.
 RX MEDLINE-99445568; Pubmed-10514494;
 RA Stephens D.J., Bantling G.;
 RT "Direct interaction of the trans-Golgi network membrane protein,
 RT TGN38, with the F-actin binding protein, neurabin.";
 RL J. Biol. Chem. 274:30080-30086(1999).
 RN [6]
 RP CHARACTERIZATION, INTERACTION WITH PPI, AND MUTAGENESIS.
 RX MEDLINE-99211883; Pubmed-10194355;

RA Hsieh-Wilson L.C., Allen P.B., Watanabe T., Nairn A.C., Greengard P.;
 RT "Characterization of the neuronal targeting protein spinophilin and
 RT its interactions with protein phosphatase-1.";
 RL Biochemistry 38:4365-4373(1999).
 CC -1- FUNCTION: BINDS TO ACTIN FILAMENTS (F-ACTIN) AND SHOWS CROSS-
 CC LINKING ACTIVITY. BINDS ALONG THE SIDES OF THE F-ACTIN. MAY PLAY
 CC AN IMPORTANT ROLE IN LINKING THE ACTIN CYTOSKELETON TO THE PLASMA
 CC MEMBRANE AT THE SYNAPTIC JUNCTION. MAY TARGET PPI TO DENDRITIC
 CC SPINES, WHICH ARE RICH IN F-ACTIN, AND REGULATE ITS SPECIFICITY
 CC TOWARD ION CHANNELS AND OTHER SUBSTRATES. MAY PLAY A ROLE IN
 CC ESTABLISHING A SIGNALING COMPLEX FOR DOPAMINERGIC
 CC NEUROTRANSMISSION THROUGH D2 RECEPTORS BY LINKING RECEPTORS
 CC DOWNSTREAM SIGNALING MOLECULES AND THE ACTIN CYTOSKELETON.
 CC -1- SUBUNIT: POSSIBLY EXISTS AS AN HOMODIMER, HOMOTRIMER OR AN
 CC HOMOTETRAMER. INTERACTS WITH F-ACTIN, PROTEIN PHOSPHATASE 1 (PPI)
 CC NEURABIN-I, TGN38 AND D(2) DOPAMINE RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: ENRICHED AT SYNAPSE AND CADHERIN-BASED CELL-
 CC CELL ADHESION SITES.
 CC -1- TISSUE SPECIFICITY: UBQUITOUS. ABUNDANTLY EXPRESSED IN THE BRAIN.
 CC EXPRESSED AT HIGHEST LEVELS IN HIPPOCAMPUS AND AT LOWER LEVELS IN
 CC THE CORTEX, CEREBELLUM AND BRAINSTEM. LOCALIZES TO THE DENDRITIC
 CC SPINES OF NEURONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS LOW DURING EMBRYOGENESIS AND
 CC INCREASES AROUND POSTNATAL DAY 21.
 CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF0216252; AAB72005.1;
 DR EMBL: AF0216252; AAC05183.1;
 DR HSSP: Q12923; 3PDZ.
 DR InterPro: IPR001478; PDZ.
 DR Pfam: PF00595; PDZ; 1.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PSS0106; PDZ; 1.
 KW Neurogenesis; Actin-binding; Cytoskeleton; Coiled coil.
 FT DOMAIN 1 154
 FT DOMAIN 2 496 584
 FT DOMAIN 3 595 616
 FT DOMAIN 4 665 816
 FT DOMAIN 5 100 371
 FT DOMAIN 6 417 494
 FT DOMAIN 7 595 816
 FT DOMAIN 8 253 260
 FT MOTAGN 447 451
 FT MOTAGN 451 451
 FT MOTAGN 451 451
 SQ SEQUENCE 817 AA: 89646 MW: D31D3E94EE2FFB6A CRC64;

Query Match 34.7%; Score 43; DB 1; Length 817;
 Best Local Similarity 39.1%; Pred. No. 47;
 Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

2 GSPAHYISARFIEGDRPSV 24
 Db 51 GSNVHRKSMELQWGTGTPGPE 73

RESULT 15
 ALLIC_BACSU STANDARD; PRT; 412 AA.
 ID ALLIC_BACSU
 AC 032149;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Allantoate amidohydrolase (EC 3.5.3.-).

GN PUCF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Bignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ehtian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghuu S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kieerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [2]
RP FUNCTION.
RX PubMed=11344136;
RA Schultz A.C., Nygaard P., Saxild H.H.;
RT "Functional analysis of 14 genes that constitute the purine catabolic
RT pathway in Bacillus subtilis and evidence for a novel regulon
RT controlled by the PucR transcription activator.";
RL J. Bacteriol. 183:3293-3302(2001).
CC -I- FUNCTION: INVOLVED IN THE ANAEROBIC UTILIZATION OF ALLANTOIN;
CC CONVERTS ALLANTOATE TO (S)-UREIDOGLYCOLATE AND AMMONIA.
CC -I- PATHWAY: Degradation of allantoin (purine catabolism); second
CC step.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M40.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z99120; CAB15243.1; -.
DR MEROPS: M40.UNM; -.
DR Subtilist: BG13994; pucf.
DR InterPro: IPR002933; Peptidase_M20.
DR Pfam: PF01546; Peptidase_M20; 1.
KW Hydrolase; Purine metabolism; Complete proteome.
SQ SEQUENCE 412 AA; 45519 MW; 852EC0952E4B5802 CRC64;

OY 1 GG---SPAHYISARFIEVG 16
II II I III :I :I
Db 381 GGVSHSPKREYTSAROLEIG 399

Search completed: May 20, 2003, 16:35:44
Job time : 8.46154 secs

Query Match 34.38; Score 42.5; DB 1; Length 412;
Best Local Similarity 52.68; Pred. No. 27;
Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 29.6703 seconds
(without alignments)
166.665 Million cell updates/sec

Title: US-09-869-155-17
Perfect score: 124
Sequence: 1 GGSFAHYISARIEVGDTRVPSVE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	39.5	218	09VE39	09ve39 drosophila
2	49	39.5	874	09FLB4	09flb4 arabidopsis
3	48	38.7	136	08YGL6	08ygl6 bruceella me
4	48	38.7	357	09PPV8	09ppv8 ureaplasma
5	48	38.7	682	067400	067400 aquifex aco
6	48	38.7	817	096SB3	096sb3 homo sapien
7	47	37.9	161	082CP1	082cp1 yersinia pe
8	46.5	37.5	250	0931Y5	0931y5 streptomyce
9	46	37.1	344	092XW9	092xw9 rhizobium m
10	46	37.1	455	0914C1	0914c1 pseudomonas
11	46	37.1	486	09KLT5	09klt5 vibrio chol
12	46	37.1	509	08UVU3	08uvu3 xenopus lae
13	46	37.1	518	08VOK9	08vok9 bruceella ab
14	46	37.1	518	08YDH3	08ydh3 bruceella me
15	46	37.1	682	042947	042947 schizosacch
16	46	37.1	714	087837	087837 streptomyce

17	46	37.1	1548	09NE01	09ne01 leishmania
18	45.5	36.7	410	09HS01	09hs01 halobacteri
19	45.5	36.7	984	097W1	097w1 mechanopyru
20	45	36.3	121	09RLA0	09rla0 streptomyce
21	45	36.3	134	09KRS3	09krs3 rhizobium l
22	45	36.3	160	067310	067310 influenza vi
23	45	36.3	370	09X517	09x517 uncultured
24	45	36.3	389	08YF5	08yff5 anabaena sp
25	45	36.3	419	0851V0	0851v0 oryza sativ
26	45	36.3	512	068181	068181 paracoccus
27	45	36.3	522	068182	068182 pseudomonas
28	45	36.3	532	096671	096671 drosophila
29	45	36.3	540	09VU53	09vu53 drosophila
30	45	36.3	625	09PJE8	09pje8 chlamydia m
31	45	36.3	852	001978	001978 caenorhabdi
32	45	36.3	955	09K6M5	09k6m5 bacillus ha
33	45	36.3	1201	016644	016644 drosophila
34	45	36.3	1201	09V4K0	09v4k0 drosophila
35	45	36.3	1348	09VAD1	09vad1 drosophila
36	45	36.3	1392	095T07	095t07 drosophila
37	45	36.3	2864	08VUE5	08vue5 erwinia chr
38	44.5	35.9	417	09Y1U6	09y1u6 pristionchu
39	44	35.5	110	047186	047186 escherichia
40	44	35.5	208	09FAT7	09fat7 paenibacill
41	44	35.5	294	092L92	092l92 rhizobium m
42	44	35.5	298	09XW86	09xw86 caenorhabdi
43	44	35.5	422	0851V3	0851v3 oryza sativ
44	44	35.5	485	09LNU3	09lnu3 arabidopsis
45	44	35.5	592	043818	043818 pisum sativ

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	218 AA.
09VE39			
ID 09VE39			
AC 09VE39			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE CG7715 protein.			
GN CG7715.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophila.			
OX NCBI_TaxID=7227;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=BERKELEY;			
RX MEDLINE=20196006; PubMed=10731132;			
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA Baason R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA Borokva D., Botchan M.R., Bouck J., Brockschtein P., Brotter P.,			
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA de Pablo R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA Fostel A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,			
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., Moleod M.P., Moberg D.,
 RA Merklov G., Mishina N.V., Moberg C., Morris J., Moshneft A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazolo M., Piltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003723; AF55590.1; -;
 DR Flybase: FBgn0038646; CG7715.
 DR InterPro: IPR002965; P_rich_extensn.
 DR PRINTS: PR01217; PRICHEXTENSN.
 SO SEQUENCE 218 AA; 23318 MW; 9F717785433971AE CRC64;

Query Match 39.5%; Score 49; DB 5; Length 218;
 Best Local Similarity 57.9%; Pred. No. 7.8;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 3 SPANISARFIEVGDTRVPSVE 21
 DB 73 SPANISARFIEVGDTRVPSVE 91

RESULT 2

O9FLB4 PRELIMINARY; PRT; 874 AA.
 AC O9FLB4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NBS/LRR disease resistance protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN 1;
 RC SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:131-145(1998).
 DR EMBL: AB010692; BAB0983.1; -;
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003591; LRR_typ.
 DR InterPro: IPR002182; NB-ARC.
 DR Pfam: PF00560; LRR; 2.
 DR Pfam: PF00931; NB-ARC; 1.
 DR PRINTS: PR00364; DISEASERSIST.
 DR SMART: SM00369; LRR_TYP; 2.
 SO SEQUENCE 874 AA; 99756 MW; 1DD55D0C163B1711 CRC64;

Query Match 39.5%; Score 49; DB 10; Length 874;
 Best Local Similarity 43.8%; Pred. No. 38;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 9 SARFIEVGDTRVPSVE 24
 DB 697 SARFIEVGDTRVPSVE 712

RESULT 3

O8YGL6 PRELIMINARY; PRT; 136 AA.
 AC O8YGL6;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Lactoylglutathione lyase (EC 4.4.1.5).
 GN BHE1142.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN 1;
 RC SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapral V., Redkar R.J., Patra G., Mufier C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jablonksi L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leccese J.J.,
 RA Haselkorn R., Kyrides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009554; AA152323.1; -;
 DR InterPro: IPR004360; Glyoxalase; 1.
 DR Pfam: PF00903; Glyoxalase; 1.
 DR Lyase; Complete proteome.
 SO SEQUENCE 136 AA; 14580 MW; 7586C1B297240749 CRC64;

Query Match 38.7%; Score 48; DB 16; Length 136;
 Best Local Similarity 38.1%; Pred. No. 6.5;
 Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 4 PAHISARFIEVGDTRVPSVE 24
 DB 39 PAHISARFIEVGDTRVPSVE 59

RESULT 4

O9PPV8 PRELIMINARY; PRT; 357 AA.
 AC O9PPV8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE XAA-PRO aminopeptidase.
 GN PEPP OR U0532.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 RN 1;
 RC SEQUENCE FROM N.A.
 RC STRAIN=SEROVAR 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glass J.I., Letkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RT "The complete sequence of the mucosal pathogen *Ureaplasma*
 RT *urealyticum*.";
 RL Nature 407:757-762(2000).
 DR EMBL: AE002152; AAP30945.1; -;
 DR HSSP: P15034; 1A16.
 DR InterPro: IPR000994; Peptidase_M24.
 DR InterPro: IPR001131; Xaa-Pro-peptids.
 DR Pfam: PF00557; Peptidase_M24; 1.
 DR PROSITE: PS00491; PROLINE-PEPTIDASE; 1.
 DR Complete proteome.
 SO SEQUENCE 357 AA; 39414 MW; EED7D034C791C2C7 CRC64;

RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Kienast H., Hopwood D.A., Denapite D., Eichner A., Cullum J.,
 RA "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 [4]

RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -!- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

CC EMBL: AL593842; CAC44213.1;
 DR InterPro: IPR000524; HTH_GNTR.
 DR Pfam: PF00392; gntr.1.
 KM DNA-binding: Transcription regulation.
 SQ SEQUENCE 250 AA; 27022 MW; CA3FE03F64F5CCD6 CRC64;

Query Match 37.5%; Score 46.5; DB 16; Length 250;
 Best Local Similarity 35.5%; Pred. No. 23;
 Matches 11; Conservative 2; Mismatches 9; Indels 9; Gaps 1;

QY 1 GSPAH-----YISARIEVGDRVPS 22
 DB 208 GGVPGHVLLVQRTFYASGRAVETADYVPA 238

RESULT 9
 092XW9 PRELIMINARY; PRT; 344 AA.
 AC 092XW9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Possible desaturase.
 GN RAL121 OR SMA2051.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barlow-Hubler F., Bowser L., Capela D., Galibert F., Gonzy J.,
 RA Gurjel M., Hong A., Hutzler L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kaimon S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSyma megaplasmid";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL: AE007298; AA65779.1;
 DR InterPro: IPR001541; Sterol_desat.
 DR Pfam: PF01598; Sterol_desat; 1.
 KM Plasmid: Complete proteome.
 SQ SEQUENCE 344 AA; 40194 MW; 1C1A72E1139F23E5 CRC64;

Query Match 37.1%; Score 46; DB 16; Length 344;

Best Local Similarity 47.4%; Pred. No. 40;
 Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 5 AHYSARIEV-GDRVP 21
 DB 282 AHYLLHKYFVNYGDTLLP 300

RESULT 10
 091AC1 PRELIMINARY; PRT; 455 AA.
 AC 091AC1;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Probable 2-isopropylmalate synthase.
 GN PA127.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004551; AAG04606.1;
 DR InterPro: IPR002034; AIPM/hcIt_synth.
 DR InterPro: IPR000891; HMGL-like.
 DR Pfam: PF00682; HMGL-like; 1.
 DR PROSITE: PS00816; AIPM_HOMOCIT_SYNTH_2; UNKNOWN_1.
 KM Complete proteome.
 SQ SEQUENCE 455 AA; 50669 MW; 3B825508D32D8FEF CRC64;

Query Match 37.1%; Score 46; DB 16; Length 455;
 Best Local Similarity 50.0%; Pred. No. 56;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 GSPAHYISARIEVGD 17
 DB 355 GSPFYIAROFHDVGD 370

RESULT 11
 09KLN5 PRELIMINARY; PRT; 486 AA.
 AC 09KLN5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pyruvate kinase II.
 GN VCA0708.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Heideberg J.F., Eisen J.A., Nelson W.C., Peterson J.D., Umayam L.A.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gili S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004400; AAF96607.1; -.
 DR HSSP: P14178; 1E0T.
 DR TIGR: VCA0708; -.
 DR InterPro: IPR001697; Pyruvate_kinase.
 DR Pfam: PF00224; PK.1.
 DR Pfam: PF02887; PK.C.1.
 DR PRINTS: PR01050; PYRUVTKNASF.
 DR PRODOM: PD001009; Pyruvate_kinase.1.
 DR TIGRFAMs: TIGR01064; pyruv_kin.1.
 DR Kinase: Complete proteome.
 SQ SEQUENCE 486 AA; 5211 MW; 3D6A239EAD495E0B CRC64;

Query Match 37.1%; Score 46; DB 16; Length 486;
 Best Local Similarity 58.3%; Pred. No. 60;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 13 IEVGDTRPSV 24
 DB 253 VEIGDARLPSVQ 264

RESULT 12

OBVU3
 ID 08VU3 PRELIMINARY; PRT; 509 AA.
 AC 08VU3;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE Suct-associated neurotrophic factor target XSNT.
 GN SNT.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Akai K., Mood K., Daar I.O.;
 RT "Fibroblast Growth Factor Receptor-Induced Mesoderm Formation in
 RT xenopus Embryos is Mediated by Adaptor Protein SNT.";
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF390895; AAL57304.1; -.
 DR InterPro: IPR002404; Insulin_receptSL.
 DR Pfam: PF02174; IRS.1.
 DR SMART: SM00310; PTB1.1.
 SQ SEQUENCE 509 AA; 56965 MW; 7FB6CEA4DCD6DF5B CRC64;

Query Match 37.1%; Score 46; DB 13; Length 509;
 Best Local Similarity 45.5%; Pred. No. 63;
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 2 GSPAHYISARFIEVGDT 23
 DB 156 GEASSHPSSRHPSVGSTRLP 177

RESULT 13

OBVOK9
 ID 08VOK9 PRELIMINARY; PRT; 518 AA.
 AC 08VOK9;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE Putative ABC transporter periplasmic binding protein.
 GN BATN1953.0RF18.
 OS Brucella abortus.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=235;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=544;
 RA Bricker B.J.;
 RT "tn1953, a new element from Brucella abortus.";
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF454951; AAL59340.1; -.
 DR InterPro: IPR000914; SBP_bac-5.
 DR Pfam: PF00496; SBP_bac.5.1.
 DR PROSITE: PS01040; SBP_BACTERIAL_5; UNKNOWN_1.
 SQ SEQUENCE 518 AA; 57014 MW; 9C2PD58CB2C03EE CRC64;

Query Match 37.1%; Score 46; DB 2; Length 518;
 Best Local Similarity 50.0%; Pred. No. 64;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 PAHYISARFIEVGDT 19
 DB 503 PAEYINIRFWEIOAK 518

RESULT 14

OBYDH3
 ID 08YDH3 PRELIMINARY; PRT; 518 AA.
 AC 08YDH3;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE Putative binding protein YDDs precursor.
 GN BME110203.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RX Delvecchio V.G., Kaprala V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykilds A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson T.-J.,
 RA Haselkorn R., Kyrides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 RL EMBL: AE009658; AAL53444.1; -.
 DR InterPro: IPR000914; SBP_bac-5.
 DR Pfam: PF00496; SBP_bac.5.1.
 DR PROSITE: PS01040; SBP_BACTERIAL_5; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 518 AA; 57091 MW; BAE42A819BD7238E CRC64;

Query Match 37.1%; Score 46; DB 16; Length 518;
 Best Local Similarity 50.0%; Pred. No. 64;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 PAHYISARFIEVGDT 19
 DB 503 PAEYINIRFWEIOAK 518

RESULT 15

042947
 ID 042947 PRELIMINARY; PRT; 682 AA.
 AC 042947;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE Hypothetical 76.7 kDa protein Cl6H5.12C in chromosome II.
 GN SPBC16H5.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022104; CAI17910.2; -;
DR InterPro: IPR004844; S/T_phosphatase.
DR Pfam: PF00149; Metallophos; 1.
KW Hypothetical protein.
SQ SEQUENCE 682 AA; 76736 MW; 015A5C53C08679C0 CRC64;

Query Match 37.1%; Score 46; DB 3; Length 682;
Best local Similarity 42.9%; Pred. No. 88;
Matches 12; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

OY 3 SPAHYIS-----ARFIEVGDTRVPSVE 24
|| ||: ||: || || ||
DB 248 SPEHYLOKLSAARAQFMEVYDTVKAIVE 275

Search completed: May 20, 2003, 16:43:18
Job time : 32.6703 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 11.0769 Seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-17

Perfect score: 124

Sequence: 1 GSPAHYISARFIEVGDTRPSVE 24

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	38.7	410	1	US-08-088-633-2
2	48	38.7	410	1	US-08-245-756-2
3	48	38.7	410	1	US-08-441-750-2
4	48	38.7	410	2	US-08-441-751-2
5	48	38.7	410	5	PCT-US92-02521-2
6	47	37.9	366	4	US-09-134-001C-5067
7	45	36.3	89	1	US-08-686-878A-21
8	45	36.3	89	4	US-09-175-928-21
9	45	36.3	1201	4	US-09-098-901-2
10	45	34.7	459	1	US-08-403-545-5
11	43	34.7	459	4	US-08-404-381-5
12	42	33.9	663	2	US-08-756-317-14
13	42	33.9	74	4	US-08-959-004-5
14	41	33.1	10	1	US-08-282-758B-31
15	41	33.1	439	4	US-09-413-814-13
16	41	33.1	508	4	US-08-980-523-9
17	40	32.3	62	4	US-08-894-626-5
18	40	32.3	74	4	US-08-995-159-5
19	40	32.3	74	4	US-08-828-683A-25
20	40	32.3	85	4	US-09-042-785A-28
21	40	32.3	201	4	US-09-064-414-2
22	40	32.3	201	4	US-09-064-414-2
23	40	32.3	208	1	US-08-618-164-3
24	40	32.3	208	4	US-09-382-155-19
25	40	32.3	208	4	US-09-074-044A-19
26	40	32.3	208	4	US-09-064-414-6
27	40	32.3	256	4	US-08-983-502-2

28	40	32.3	256	5	PCT-US95-16542-2	Sequence 2, Appl1
29	40	32.3	256	5	PCT-US96-10521-2	Sequence 2, Appl1
30	40	32.3	760	4	US-09-513-057C-27	Sequence 27, Appl1
31	40	32.3	1802	4	US-09-322-478-18	Sequence 18, Appl1
32	39	31.5	405	4	US-09-134-001C-5259	Sequence 5259, Ap
33	39	31.5	442	4	US-09-357-251-18	Sequence 18, Appl1
34	39	31.5	445	4	US-09-810-671-2	Sequence 2, Appl1
35	38	30.6	162	1	US-07-961-702-2	Sequence 2, Appl1
36	38	30.6	162	1	US-08-472-284-2	Sequence 2, Appl1
37	38	30.6	162	1	US-08-476-678-2	Sequence 2, Appl1
38	38	30.6	162	2	US-08-472-418-2	Sequence 2, Appl1
39	38	30.6	287	4	US-09-371-056-14	Sequence 14, Appl1
40	38	30.6	349	1	US-08-118-270-71	Sequence 71, Appl1
41	38	30.6	349	5	PCT-US93-08528-71	Sequence 81, Appl1
42	38	30.6	389	4	US-09-594-193-8	Sequence 8, Appl1
43	38	30.6	553	2	US-08-586-272-2	Sequence 2, Appl1
44	38	30.6	553	3	US-09-082-969-2	Sequence 2, Appl1
45	38	30.6	612	2	US-08-359-705B-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-088-633-2
; Sequence 2, Application US/08088633
; Patent No. 5324660
; GENERAL INFORMATION:
; APPLICANT: Gleeson, Martin A
; APPLICANT: Howard, Bradley D
; TITLE OF INVENTION: Genes which influence Pichia proteolytic
; TITLE OF INVENTION: Genes which influence Pichia proteolytic
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fitch, Even, Tabin & Plannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/088,633
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/678,916
; FILING DATE: 01-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: 50848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; TELEX: 20 6566 PATLAW CGO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-088-633-2
Query Match 38.7%; Score 48; DB 1; Length 410;
Best Local Similarity. 34.8%; Pred. No. 3.6;
Matches 8; Conservative 6; Mismatches 9; Gaps 0;
QY 2 GSPAHYISARFIEVGDTRPSVE 24

Db 160 GSMGYSQDVLDIGDLTIPKVD 182

RESULT 2

US-08-245-756-2
Sequence 2, Application US/08245756
Patent No. 5541112
GENERAL INFORMATION:
APPLICANT: Gleeson, Martin A
APPLICANT: Howard, Bradley D
TITLE OF INVENTION: GENES WHICH INFLUENCE PICHA PROTEOLYTIC
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
STREET: 444 So. Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,756
FILING DATE: 16-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/088,633
FILING DATE: 06-JULY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/678,916
FILING DATE: 01-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 9763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-245-756-2

Query Match 38.7%; Score 48; DB 1; Length 410;
Best Local Similarity 34.8%; Pred. No. 3.6;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 2 GSPAHYISARFIEVGDTGPSVE 24

Db 160 GSMGYSQDVLDIGDLTIPKVD 182

RESULT 3

US-08-441-750-2
Sequence 2, Application US/08441750
Patent No. 5691166
GENERAL INFORMATION:
APPLICANT: Gleeson, Martin A
APPLICANT: Howard, Bradley D
TITLE OF INVENTION: GENES WHICH INFLUENCE PICHA PROTEOLYTIC
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street

CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,750
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/245,756
FILING DATE: 16-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/088,633
FILING DATE: 06-JULY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/678,916
FILING DATE: 01-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 9763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-441-750-2

Query Match 38.7%; Score 48; DB 1; Length 410;
Best Local Similarity 34.8%; Pred. No. 3.6;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 2 GSPAHYISARFIEVGDTGPSVE 24

Db 160 GSMGYSQDVLDIGDLTIPKVD 182

RESULT 4

US-08-441-751-2
Sequence 2, Application US/08441751
Patent No. 5831053
GENERAL INFORMATION:
APPLICANT: Gleeson, Martin A
APPLICANT: Howard, Bradley D
TITLE OF INVENTION: GENES WHICH INFLUENCE PICHA PROTEOLYTIC
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,751
FILING DATE:
CLASSIFICATION: 435

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/245,756
; FILING DATE: 16-May-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/088,633
; FILING DATE: 06-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/678,916
; FILING DATE: 01-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 9763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-751-2

Query Match
Best local Similarity 38.7%; Score 48; DB 2; Length 410;
Pred. No. 3.6;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 2 GSPAHYISARFIEVGDTRPSVE 24
  1 1 1 : : : : : 1 1 :
Db 160 GSMEGYVSQDVLQIGDLTIKVD 182

RESULT 5
PCT-US92-02521-2
; Sequence 2, Application PC/TUS920202521
; GENERAL INFORMATION:
; APPLICANT: Gleeson, Martin A
; APPLICANT: Howard, Bradley D
; TITLE OF INVENTION: GENES WHICH INFLUENCE PICHIA PROTEOLYTIC
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02521
; FILING DATE: 19920321
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/678,916
; FILING DATE: 01-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 50848PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; TELEX: 20 6566 PATLAW CGO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids

```

```

;
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02521-2

Query Match
Best local Similarity 38.7%; Score 48; DB 5; Length 410;
Pred. No. 3.6;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 2 GSPAHYISARFIEVGDTRPSVE 24
  1 1 1 : : : : : 1 1 :
Db 160 GSMEGYVSQDVLQIGDLTIKVD 182

RESULT 6
US-09-134-001C-5067
; Sequence 5067, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et. al
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5067
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5067

Query Match
Best local Similarity 37.9%; Score 47; DB 4; Length 366;
Pred. No. 4.7;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GSPAHYISARFIEVGDTRVP 21
  1 1 : : : : : 1 1 :
Db 149 GASSEVEVRNNGVEGDTIYV 169

RESULT 7
US-08-686-878A-21
; Sequence 21, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racine, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; TITLE OF INVENTION: SPULIDING, VIKKI
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/686,878A
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brown, Scott A.
;; REGISTRATION NUMBER: 32,724
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 498-8224
;; TELEFAX: (617) 876-5851
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 89 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-686-878A-21

Query Match 36.3%; Score 45; DB 4; Length 89;
Best Local Similarity 72.7%; Pred. No. 1.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSPAHYISARF 12
1:1111111111
Db 34 GTPAGYVSARF 44

RESULT 8
US-09-175-928-21
;; Sequence 21, Application US/09175928A
;; Patent No. 6312921
;; GENERAL INFORMATION:
;; APPLICANT: Jacobs, Kenneth
;; APPLICANT: McCoy, John M.
;; APPLICANT: Lavaille, Edward R.
;; APPLICANT: Collins-Racle, Lisa A.
;; APPLICANT: Evans, Cheryl
;; APPLICANT: Merberg, David
;; APPLICANT: Treacy, Maurice
;; APPLICANT: Mt. Sha
;; APPLICANT: Genetics Institute, Inc.
;; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;; FILE REFERENCE: 6006B.A1172A
;; CURRENT APPLICATION NUMBER: US/09/175,928A
;; CURRENT FILING DATE: 1998-10-20
;; NUMBER OF SEQ ID NOS: 62
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 21
;; LENGTH: 89
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: (18)
US-09-175-928-21

Query Match 36.3%; Score 45; DB 4; Length 89;
Best Local Similarity 72.7%; Pred. No. 1.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSPAHYISARF 12
1:1111111111
Db 34 GTPAGYVSARF 44

RESULT 9
US-09-098-901-2
;; Sequence 2, Application US/09098901B
;; Patent No. 6218144
;; GENERAL INFORMATION:
;; APPLICANT: Scott, Matthew
;; APPLICANT: Sisson, John C.

;; TITLE OF INVENTION: Costal2 Genes and their Uses
;; FILE REFERENCE: SUN-65P
;; CURRENT APPLICATION NUMBER: US/09/098,901B
;; CURRENT FILING DATE: 1998-06-17
;; EARLIER APPLICATION NUMBER: 60/051,347
;; EARLIER FILING DATE: 1997-06-30
;; NUMBER OF SEQ ID NOS: 11
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 1201
;; TYPE: PRT
;; ORGANISM: D. Melanogaster
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (1)...(1201)
;; OTHER INFORMATION: xaa = Any Amino Acid
US-09-098-901-2

Query Match 36.3%; Score 45; DB 4; Length 1201;
Best Local Similarity 47.6%; Pred. No. 4.5;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 GSPAHYISARFIEVGTVP 21
111111111111
Db 117 GSDPMSIAAGLIQVGPVTP 137

RESULT 10
US-08-403-545-5
;; Sequence 5, Application US/08403545
;; Patent No. 5656483
;; GENERAL INFORMATION:
;; APPLICANT: Sokatch, John R.
;; APPLICANT: Sykes, Pamela Joy
;; APPLICANT: Madhusudhan, K.T.
;; TITLE OF INVENTION: Genes Encoding Operon and Promoter for
;; TITLE OF INVENTION: Branch Chain Keto Acid Dehydrogenase of Pseudomonas putid
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Carolyn D. Moon
;; ADDRESSEE: Dunlap, Coddling, Perterson and Lee
;; STREET: 9400 N. Broadway, Suite 420
;; CITY: Oklahoma City
;; STATE: Oklahoma
;; COUNTRY: USA
;; ZIP: 73114
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: diskette 5.25 inch, 360 Kb Storage
;; COMPUTER: IBM AT
;; OPERATING SYSTEM: MS-DOS Version 3.3
;; SOFTWARE: Professional Write 2.2
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/403,545
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 7/603/781
;; APPLICATION NUMBER: 07/172,148
;; FILING DATE: 23-003-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carolyn D. Moon
;; REGISTRATION NUMBER: 33,022
;; REFERENCE/DOCKET NUMBER: 5820.101
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: Attorney, (405) 478-5344
;; TELEFAX: Attorney, (405) 478-5349
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 459 Amino acids
;; TYPE: Amino acid
;; TOPOLOGY: linear

MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Lpd-val, the E3 component
LOCATION: 4369-5745, N-terminal methionine is present on
LOCATION: mature protein
IDENTIFICATION METHOD: Sequence of cyanogen bromide peptides
Patent No. 5656483
PUBLICATION INFORMATION:
AUTHORS: Burns, Gayle
AUTHORS: Brown, Tracy
AUTHORS: Hatter, Kenneth
AUTHORS: Sokatch, John R.
TITLE: Sequence analysis of the lpdv gene for lipoamide
Patent No. 5656483
TITLE: dehydrogenase of Pseudomonas putida
JOURNAL: European Journal of Biochemistry
VOLUME: 179
PAGES: 61-69
DATE: 1989
US-08-403-545-5

Query Match 34.7%; Score 43; DB 1; Length 459;
Best Local Similarity 39.1%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

OY 1 GGSPAHYISARFIEVGDTRVPSV 23
|||::|||::|||
Db 13 GGGPGGYVAA--IRAGQLGIPV 33

RESULT 11

US-08-404-381-5
Sequence 5, Application US/08404381
Patent No. 6168945

GENERAL INFORMATION:

APPLICANT: Sokatch, John R.
APPLICANT: Sykes, Pamela Joy
APPLICANT: Madhusudhan, K. T.
TITLE OF INVENTION: Genes Encoding Operon and Promoter for
TITLE OF INVENTION: Branched Chain Keto Acid Dehydrogenase of Pseudomonas putida
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carolyn D. Moon
ADDRESS: 9400 N. Broadway, Suite 420
CITY: Oklahoma City
STATE: Oklahoma
COUNTRY: USA
ZIP: 73114

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 5.25 inch, 360 kb Storage
COMPUTER: IBM AT
OPERATING SYSTEM: MS-DOS Version 3.3
SOFTWARE: Professional Write 2.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,381
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,781
FILING DATE:

APPLICATION NUMBER: 07/172,148
FILING DATE: 23-003-1988

ATTORNEY/AGENT INFORMATION:
NAME: Carolyn D. Moon

REGISTRATION NUMBER: 33,022
REFERENCE/DOCKET NUMBER: 5820.101

TELECOMMUNICATION INFORMATION:
TELEPHONE: Attorney, (405) 478-5344

TELEFAX: Attorney, (405) 478-5349
TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 459 Amino acids
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Protein
FEATURE:

NAME/KEY: Lpd-val, the E3 component
LOCATION: 4369-5745, N-terminal methionine is present on
LOCATION: mature protein
IDENTIFICATION METHOD: Sequence of cyanogen bromide peptides
Patent No. 6168945

PUBLICATION INFORMATION:

AUTHORS: Burns, Gayle
AUTHORS: Brown, Tracy
AUTHORS: Hatter, Kenneth
AUTHORS: Sokatch, John R.

TITLE: Sequence analysis of the lpdv gene for lipoamide
Patent No. 6168945
TITLE: dehydrogenase of Pseudomonas putida
JOURNAL: European Journal of Biochemistry
VOLUME: 179
PAGES: 61-69
DATE: 1989

US-08-404-381-5
Query Match 34.7%; Score 43; DB 4; Length 459;
Best Local Similarity 39.1%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

OY 1 GGSPAHYISARFIEVGDTRVPSV 23
|||::|||::|||
Db 13 GGGPGGYVAA--IRAGQLGIPV 33

RESULT 12

US-08-756-317-14
Sequence 14, Application US/08756317
Patent No. 5849894

GENERAL INFORMATION:

APPLICANT: Clemente, Thomas E.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Milsky, Timothy A.
APPLICANT: Stark, David M.
TITLE OF INVENTION: Improved Rhodospirillum Rubrum
TITLE OF INVENTION: Poly-B-Hydroxyalkanoate Synthase
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,317
FILING DATE: 25-NOV-1996

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,693
FILING DATE: 29-NOV-1995

ATTORNEY/AGENT INFORMATION:
NAME: Patterson, Melinda L.

REGISTRATION NUMBER: 33,062
REFERENCE/DOCKET NUMBER: MOBT:008

TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 787-1400

TELEFAX: (713) 787-1440
TELEX:

INFORMATION FOR SEQ ID NO: 14:


```

? APPLICANT: Reichenbach, Hans
? TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
? TITLE OF INVENTION: heteropolyleptide compounds
? FILE REFERENCE: PCT/US 99/23535
? CURRENT APPLICATION NUMBER: US/09/413,814
? CURRENT FILING DATE: 1999-10-07
? EARLIER APPLICATION NUMBER: DE 198 46 493.2
? EARLIER FILING DATE: 1998-10-09
? NUMBER OF SEQ ID NOS: 107
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO: 13
? LENGTH: 439
? TYPE: PRT
? ORGANISM: Sorangium cellulosum
? US-09-413-814-13

```

Query Match	33.1%;	Score 41;	DB 4;	Length 439;
Best Local Similarity	41.2%;	Pred. No. 62;		
Matches	7;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      8 ISARFIEVGDTRVPSVE 24
          : || : || ||
Db     36 VEERFFAAGEDRLPEVE 52
```

Search completed: May 20, 2003, 16:44:43
Job time : 13.0769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 ; Search time 11,8681 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-17

Perfect score: 124
Sequence: 1 GSPAHYISARFIEVGDTRVPSVE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCRT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCRTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	36.3	89	US-10-114-893-150	Sequence 150, App
2	45	36.3	89	US-10-040-916-21	Sequence 21, App
3	43.5	35.1	70	US-09-864-761-47948	Sequence 47948, A
4	43	34.7	190	US-09-864-761-33402	Sequence 33402, A
5	43	34.7	190	US-09-815-242-11866	Sequence 11866, A
6	41	33.1	390	US-09-815-242-11173	Sequence 11173, A
7	41	33.1	448	US-09-815-242-11160	Sequence 11160, A
8	41	33.1	469	US-09-815-242-10418	Sequence 10418, A
9	41	33.1	469	US-09-815-242-13789	Sequence 13789, A
10	41	33.1	508	US-09-757-415A-1	Sequence 1, Appl
11	41	33.1	538	US-09-758-269-8	Sequence 8, Appl
12	41	33.1	538	US-09-758-269-18	Sequence 18, Appl
13	41	33.1	538	US-09-758-269-33	Sequence 33, Appl
14	41	33.1	915	US-10-029-180-74	Sequence 74, Appl
15	40.5	32.7	607	US-09-344-882-18	Sequence 18, Appl
16	40	32.3	40	US-09-764-891-5300	Sequence 5300, Ap
17	40	32.3	62	US-10-035-408-5	Sequence 5, Appl
18	40	32.3	74	US-10-112-793-25	Sequence 25, Appl
19	40	32.3	176	US-10-168-066-6	Sequence 6, Appl

20	40	32.3	256	10	US-09-933-814-2	Sequence 2, Appl
21	40	32.3	256	10	US-09-824-134-2	Sequence 2, Appl
22	40	32.3	383	9	US-09-738-626-4233	Sequence 4233, Ap
23	40	32.3	394	9	US-09-738-626-4273	Sequence 4273, Ap
24	40	32.3	760	10	US-09-746-801A-27	Sequence 27, Appl
25	40	32.3	842	10	US-09-798-831-8	Sequence 8, Appl
26	40	32.3	1802	10	US-09-965-553-18	Sequence 18, Appl
27	39.5	31.9	765	9	US-10-078-770-84	Sequence 94, Appl
28	39.5	31.9	818	9	US-10-078-770-94	Sequence 94, Appl
29	39	31.5	159	9	US-09-847-208-35	Sequence 35, Appl
30	39	31.5	159	9	US-09-847-208-41	Sequence 41, Appl
31	39	31.5	419	9	US-09-738-626-6634	Sequence 6634, Appl
32	39	31.5	445	12	US-10-109-854-2	Sequence 2, Appl
33	38.5	31.0	8	9	US-09-764-891-33657	Sequence 3657, Ap
34	38.5	31.0	106	9	US-09-764-891-33990	Sequence 3990, Ap
35	38.5	31.0	411	9	US-10-125-540-351	Sequence 351, App
36	38.5	31.0	411	10	US-09-764-870-351	Sequence 351, App
37	38	30.6	66	9	US-09-764-891-5186	Sequence 5186, Ap
38	38	30.6	68	9	US-09-764-891-5313	Sequence 5313, Ap
39	38	30.6	91	10	US-09-864-761-36795	Sequence 36795, A
40	38	30.6	178	9	US-09-798-889-76	Sequence 96, Appl
41	38	30.6	238	9	US-10-078-770-96	Sequence 96, Appl
42	38	30.6	265	10	US-09-815-242-5157	Sequence 5157, Ap
43	38	30.6	287	12	US-10-081-301-14	Sequence 14, Appl
44	38	30.6	298	10	US-09-815-242-5142	Sequence 5142, Ap
45	38	30.6	305	9	US-10-078-770-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-10-114-893-150
; Sequence 150, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racle, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893.
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 150
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (18)
US-10-114-893-150

Query Match 36.3%; Score 45; DB 9; Length 89;
Best Local Similarity 72.7%; Pred. No. 7.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GSPAHYISARF 12
I: I I I I I I I I I I
DB 34 GTPAGYVSARF 44

RESULT 2
US-10-040-916-21
Sequence 21, Application US/10040916
Patent NO. US20020146769A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John
Lavallee, Edward
Racie, Lisa
Merberg, David
Treacy, Maurice
Evans, Cheryl
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,916
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/887,029
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 08/686,878
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-040-916-21

Query Match 36.3%; Score 45; DB 12; Length 89;
Best Local Similarity 72.7%; Pred. No. 7.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSPAHYSARF 12
I:1111111111
DB 34 GTPAGTYSARF 44

RESULT 3
US-09-864-47948
Sequence 47948, Application US/09864761
Patent NO. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aomicla-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 47948
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006042.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
OTHER INFORMATION: EST_HUMAN HIT: R45715.1, EVALU 9.00e-19
US-09-864-761-47948

Query Match 35.1%; Score 43.5; DB 10; Length 70;
Best Local Similarity 47.6%; Pred. No. 9.4;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 4 PAHYISARFIEVGDTVPVSVE 24
IIIIII:II:IIII
DB 13 PAHYRSSSTRSI-DTQPSVO 32

RESULT 4
US-09-864-761-33402
Sequence 33402, Application US/09864761
Patent NO. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33402
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006333.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.5
; OTHER INFORMATION: EST_HUMAN HIT: BE466332.1, EVALUE 8.00e-36
; OTHER INFORMATION: SWISSPROT HIT: Q02357, EVALUE 6.00e-08
; US-09-864-761-33402

Query Match          34.7%  Score 43:  DB 10:  Length 190:
Best Local Similarity 50.0%  Pred. No. 34:
Matches      8:  Conservative      3:  Mismatches      5:  Indels      0:  Gaps      0:
```

```
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11866
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-815-242-11866

Query Match          34.7%  Score 43:  DB 10:  Length 464:
Best Local Similarity 39.1%  Pred. No. 93:
Matches      9:  Conservative      4:  Mismatches      8:  Indels      2:  Gaps      1:

Db      13  GGGPGGYVAA--IRAGQLGIPY 33

RESULT 6
US-09-815-242-11173
; Sequence 11173, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11173
; LENGTH: 390
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-09-815-242-11173
```

```
Query Match          33.1%; Score 41; DB 10; Length 390;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      5 AHYISARFIEVGDRV 20
        |||:|||||:|:
DB      30 AHYNAFIEIGAGPRM 45
```

```
RESULT 7
US-09-815-242-11160
; Sequence 11160, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11160
; LENGTH: 448
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-09-815-242-11160
```

```
Query Match          33.1%; Score 41; DB 10; Length 448;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4 PAHYIS-ARFIEVGDR 19
        |||:|:|:|:|:|:
DB      378 PPHYDSMIKILTYGDR 395
```

```
RESULT 8
US-09-815-242-10418
```

```

; Sequence 10418, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10418
; LENGTH: 469
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10418
```

```
Query Match          33.1%; Score 41; DB 10; Length 469;
Best Local Similarity 53.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      4 PAHYISARFIEVG 16
        |||:|:|:|:|:
DB      35 PAHVNAEFEEEG 47
```

```
RESULT 9
US-09-815-242-13789
; Sequence 13789, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13789
LENGTH: 469
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13789

Query Match 33.1%; Score 41; DB 10; Length 469;
Best Local Similarity 53.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 PAHYISARFIEVGDTRVPSV 16
| | | : | | | | |
DB 35 PAHOVNAEFFEEG 47

RESULT 10
US-09-757-415A-1
Sequence 1, Application US/09757415A
Publication No. US20030040612A1
GENERAL INFORMATION:
APPLICANT: Zhou, Ming-Ming
TITLE OF INVENTION: Methods of Identifying Modulators of the FGF Receptor
FILE REFERENCE: 2459-1-002N
CURRENT APPLICATION NUMBER: US/09/757,415A
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/175867
PRIOR FILING DATE: 2000-01-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 508
TYPE: PRT
ORGANISM: Homo sapien
US-09-757-415A-1

Query Match 33.1%; Score 41; DB 9; Length 508;
Best Local Similarity 40.9%; Pred. No. 2.1e+02;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 GSPAHYISARFIEVGDTRVPSV 23
| | : | | | | | | | | |
DB 157 GDASHPSRHPVSGARLPV 178

RESULT 11
US-09-758-269-8
Sequence 8, Application US/09758269
Patent No. US20020104120A1
GENERAL INFORMATION:
APPLICANT: IUCHI, SATOSHI
APPLICANT: KOBAYASHI, MASATOMO
APPLICANT: SHINOZAKI, KAZUO
TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
FILE REFERENCE: 3914-3
CURRENT APPLICATION NUMBER: US/09/758,269
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: JP 2001-003476
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: JP 2000-010056
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 538

TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-758-269-8

Query Match 33.1%; Score 41; DB 10; Length 538;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 PAHYISARFIEVGDTRVP 21
| | | : | | | | |
DB 44 PLHYLSGNFAPIRDTPP 61

RESULT 12
US-09-758-269-18
Sequence 18, Application US/09758269
Patent No. US20020104120A1
GENERAL INFORMATION:
APPLICANT: IUCHI, SATOSHI
APPLICANT: KOBAYASHI, MASATOMO
APPLICANT: SHINOZAKI, KAZUO
TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
FILE REFERENCE: 3914-3
CURRENT APPLICATION NUMBER: US/09/758,269
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: JP 2001-003476
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: JP 2000-010056
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 538
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-758-269-18

Query Match 33.1%; Score 41; DB 10; Length 538;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 PAHYISARFIEVGDTRVP 21
| | | : | | | | |
DB 44 PLHYLSGNFAPIRDTPP 61

RESULT 13
US-09-758-269-33
Sequence 33, Application US/09758269
Patent No. US20020104120A1
GENERAL INFORMATION:
APPLICANT: IUCHI, SATOSHI
APPLICANT: KOBAYASHI, MASATOMO
APPLICANT: SHINOZAKI, KAZUO
TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
FILE REFERENCE: 3914-3
CURRENT APPLICATION NUMBER: US/09/758,269
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: JP 2001-003476
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: JP 2000-010056
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
LENGTH: 538
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-758-269-33

Query Match 33.1%; Score 41; DB 10; Length 538;

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 ; Search time 14.1429 Seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-18

Perfect score: 65

Sequence: 1 VNVGVLAACAPSK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq.101002:*
1: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	100.0	13	21	AAV93761
2	65	100.0	13	22	AAE14692
3	65	100.0	13	22	AAU07397
4	65	100.0	370	23	AAU75815
5	65	100.0	370	23	AAU75816
6	65	100.0	381	23	AAU75812
7	65	100.0	381	23	AAU75813
8	65	100.0	381	23	AAU75814
9	50	76.9	287	23	AAU75819
10	43	66.2	277	23	AAU75818

11	41	63.1	25	20	AAW81605	Synthetic rabies n
12	39	60.0	25	9	AAE08689	Sequence corresp.
13	39	60.0	25	10	AAE91478	Peptide N-V12b. R
14	39	60.0	25	20	AAW81603	Synthetic peptide
15	39	60.0	341	21	AAE42355	Arabidopsis thalia
16	39	60.0	342	21	AAE42354	Arabidopsis thalia
17	39	60.0	450	8	AAE70317	Protein N encoded
18	38	58.5	26	20	AAW81612	Synthetic peptide
19	38	58.5	35	22	AAE31621	Peptide #4272 enco
20	38	58.5	35	22	AAE36838	Peptide #4344 enco
21	38	58.5	35	22	AAE22161	Protein #4160 enco
22	38	58.5	35	22	AAE57582	Human brain expres
23	38	58.5	35	22	AAE69586	Human bone marrow
24	38	58.5	35	22	AAE17817	Peptide #4251 enco
25	38	58.5	35	22	AAE03324	Peptide #4361 enco
26	38	58.5	35	22	AAE05464	Peptide #4146 enco
27	38	58.5	35	23	AAE39615	Human peptide enco
28	38	58.5	63	22	AAE66672	Novel human connec
29	38	58.5	197	21	AAE41482	Human OREX ORF1246
30	38	58.5	197	23	AAE32724	Human OREX ORF1246
31	37	56.9	116	22	AAE11090	Human hydrophobic
32	37	56.9	529	22	AAE44300	Propionibacterium
33	36	55.4	54	22	AAE24978	Novel human diagno
34	36	55.4	89	20	AAE35915	Extended human sec
35	36	55.4	159	22	AAE15894	Novel human diagno
36	36	55.4	333	21	AAE58749	Arabidopsis thalia
37	36	55.4	361	22	AAE04048	Novel human diagno
38	36	55.4	369	21	AAE52007	Arabidopsis thalia
39	36	55.4	377	21	AAE53880	A rabbit CD14 prot
40	36	55.4	390	20	AAE37383	Amino acid sequenc
41	36	55.4	445	21	AAE17005	Arabidopsis thalia
42	36	55.4	445	21	AAE52006	Arabidopsis thalia
43	36	55.4	459	20	AAE38604	Neisseria meningit
44	36	55.4	478	22	AAE30648	Novel human secret
45	36	55.4	486	22	AAE14794	Novel human diagno

ALIGNMENTS

RESULT 1
ID AAV93761 standard; peptide: 13 AA.
XX
AC AAV93761;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.
XX
KW Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
XX
XX dough; dough preparation.
XX
OS Trillium sp.
XX
PN WO200039289-A2;
XX
PD 06-JUL-2000.
XX
PE 17-DEC-1999; 99WO-IB02071.
XX
PR 23-DEC-1998; 98GB-0028599.
XX
PR 06-APR-1999; 99GB-0007805.
XX
PR 15-APR-1999; 99GB-0008645.
XX
PA (DANT-) DANISCO AS.
XX
PI Sibbesen O, Sorensen JF;
XX
XX WPI: 2000-465744/40..
XX
PT Mutant xylanase protein identified using xylanase inhibitor useful for
PT preparing non-sticky dough for bakery products .

XX Claim 24; Page 112; 112pp; English.
PS
XX
CC The present sequence is derived from an endo-beta-1,4-xylanase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xylanase protein. The xylanase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xylanase or
CC mutant xylanase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xylanase. The xylanase inhibitor is
CC useful for screening high degree resistance xylanases for dough
CC preparation. The xylanase is also useful for preparing a non-sticky
CC dough. A combination of xylanase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 65; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VNWGVLAACAPSK 13
Db 1 VNWGVLAACAPSK 13
RESULT 2
AEL14692
ID AEL14692 standard; peptide: 13 AA.
XX
AC AEL14692;
XX
DT 21-AUG-2002 (first entry)
XX
DE wheat flour xylanase inhibitor B chain Lys-C digested fragment #3.
XX
KM Refrigerated dough; syruping; arabinoxylan; bakery product; bread;
KM pizza base; cake; biscuit; wheat; flour; xylanase inhibitor.
XX
OS Triticum aestivum.
XX
PN WO200152657-A1.
XX
PD 26-JUL-2001.
XX
PF 17-JAN-2001; 2001WO-IB00168.
XX
PR 18-JAN-2000; 2000GB-0001136.
XX
PA (DANI-) DANISCO AS.
XX
PI Sorensen CH, Sorensen JF;
XX
DR WPI; 2001-457446/49.
XX
PT Production of refrigerated dough with reduced syruping, useful in
PT production of bakery products such as bread, comprises admixing cereal
PT flour, water and protein that prevents enzymatic degradation of
PT arabinoxylan in the cereal flour -
XX
PS Disclosure; Page 23; 26pp; English.
XX
CC The invention relates to a process for producing refrigerated dough
CC with reduced 'syruping' (precipitation of liquid on the dough surface
CC because of a reduction in water holding capacity caused by the breakdown
CC of arabinoxylan over time). The process comprises admixing cereal flour
CC and water with a protein that reduces/prevents enzymatic degradation of
CC arabinoxylan in the cereal flour. The preferred protein is a xylanase
CC inhibitor. The method is useful to produce refrigerated dough in which
CC syruping is reduced or eliminated. Refrigerated dough is typically
CC stored for long periods to enable fresh baked products (e.g. bread,
CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
CC requirements using the method by the use of specific proteins/protein
CC combinations. The present sequence is wheat flour
CC endo-beta-1,4-xylanase inhibitor B chain Lys-C digested fragment.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 65; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VNWGVLAACAPSK 13
Db 1 VNWGVLAACAPSK 13
RESULT 3
AAU07397
ID AAU07397 standard; protein: 13 AA.
XX
AC AAU07397;
XX
DT 18-DEC-2001 (first entry)
XX
DE Bacillus subtilis xylanase inhibitor #6.
XX
KM Xylanase; plant cell wall; baking; cereal; starch production; wood;
KM wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
XX
OS Bacillus subtilis.
XX
PN WO200166711-A1.
XX
PD 13-SEP-2001.
XX
PF 08-MAR-2001; 2001WO-IB00426.
XX
PR 08-MAR-2000; 2000GB-0005585.
XX
PR 27-JUN-2000; 2000GB-0015751.
XX
PA (DANI-) DANISCO AS.
XX
PI Sibbesen O, Sorensen JF;
XX
DR WPI; 2001-596834/67.
XX
PT Novel variant xylanase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xylanase inhibitor -
XX
PS Disclosure; Page 63; 70pp; English.
XX
CC The invention relates to a variant xylanase polypeptide (I) or its
CC fragment having xylanase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xylanase inhibitor as compared with the parent xylanase enzyme. (I) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I)
CC is useful for altering the viscosity derived from the presence of
CC hemicellulose or arabinoxylan in a solution or system comprising plant
CC cell wall material. (I) is useful for preparing a foodstuff such as
CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
CC The present sequence represents the amino acid sequence of Bacillus
CC subtilis xylanase inhibitor #6 as described in the method of the
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 65; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNWGYLAACAPSK 13
 |||||||||

Db 1 VNWGYLAACAPSK 13

RESULT 4
 ID AAU75815 standard; Protein; 370 AA.
 AC AAU75815;
 DT 23-APR-2002 (first entry)

Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.

Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 noodle; animal feed; starch separation; maize processing; malting;
 plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

Triticum aestivum cultivar Soissons.
 WO200198474-A1.
 27-DEC-2001.

21-JUN-2001; 2001WO-BE00106.

22-JUN-2000; 2000GB-0015296.
 25-JAN-2001; 2001GB-0002018.
 26-JAN-2001; 2001GB-0002194.
 16-MAR-2001; 2001GB-0006564.
 21-MAY-2001; 2001GB-0012328.

(LEUV-) LEUVEN RES & DEV.

Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 Van Campenhout S;
 WPI; 2002-114579/15.
 N-PSDB; ABK13672.

Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 beta-glucanolytic enzymes comprising using endoxylanases during
 screening for inhibition activity or affinity chromatography with
 immobilised enzymes

Claim 127; Page 57; 127pp; English.

The invention relates to separating and/or isolating inhibitors of
 cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 screening the inhibition activity by using two or more enzymes during the
 separation and/or isolation steps that allow to distinguish inhibitors of
 different specificity or by using an affinity chromatographic step with
 immobilised enzymes and/or antibodies against inhibitors. Also
 included are an isolated nucleic acid molecule encoding an inhibitor
 which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 nucleic acid, an expression system transformed with the nucleic
 acid, a host organism transformed with the nucleic acid, the inhibitory
 proteins encoded by the nucleic acids and modulators of the proteins.
 A recombinant protein, glycoprotein or polypeptide or microorganisms,
 plant or plant materials transformed with the nucleic acid are useful
 for the formation of an endoxylanase-inhibitor complex, screening
 endoxylanases that are totally, less or not inhibited by the inhibitors,
 reducing syruiping in refrigerated dough compositions, affecting the
 relative affinity and/or relative hydrolysis specificity and/or relative
 hydrolysis rate versus water-extractable and/or water-unextractable

arabinoxylans of endoxylanases such as by the formation of an
 endoxylanase/inhibitor complex, improving the malting of cereals such as
 barley, sorghum and wheat and/or the production of beer, improving the
 production and/or quality of baked or extruded cereal products such as
 straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 biscuits, pasta and noodles, animal feed stuff, improving the production
 of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 starch separation and production, improving maize processing,
 plant disease resistance and nutraceutical and/or pharmaceutical
 applications, improving paper and pulp technologies. The present
 sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase
 inhibitor).

Sequence 370 AA;

Query Match 100.0%; Score 65; DB 23; Length 370;
 Best Local Similarity 100.0%; Pred. No. 0.00093;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNWGYLAACAPSK 13
 |||||||||

Db 105 VNWGYLAACAPSK 117

RESULT 5
 ID AAU75816 standard; Protein; 370 AA.
 AC AAU75816;
 DT 23-APR-2002 (first entry)

Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.02.

Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 immobilised enzyme; enzyme; dough syruiping; cereal product; beer; plant;
 straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 noodle; animal feed; starch separation; maize processing; malting;
 plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

Triticum aestivum cultivar Estica.
 WO200198474-A1.
 27-DEC-2001.

21-JUN-2001; 2001WO-BE00106.

22-JUN-2000; 2000GB-0015296.
 25-JAN-2001; 2001GB-0002018.
 26-JAN-2001; 2001GB-0002194.
 16-MAR-2001; 2001GB-0006564.
 21-MAY-2001; 2001GB-0012328.

(LEUV-) LEUVEN RES & DEV.

Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 Van Campenhout S;
 WPI; 2002-114579/15.
 N-PSDB; ABK13673.

Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 beta-glucanolytic enzymes comprising using endoxylanases during
 screening for inhibition activity or affinity chromatography with
 immobilised enzymes

Claim 127; Page 57; 127pp; English.

The invention relates to separating and/or isolating inhibitors of
 cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 screening the inhibition activity by using two or more enzymes during the

separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulose, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins, A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syringing in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylanases such as by the formation of an endoxylanase/inhibitor complex, improving the malting of cereals such as barley, sorghum and wheat and/or the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten starch separation and production, improving maize processing, plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence represents a partial wheat TAXI-I (T. aestivum L endoxylanase inhibitor).

SO Sequence 370 AA:

Query Match 100.0%; Score 65; DB 23; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNWGYLAACPSK 13
IIIIIIIIIIIIIIIIIIII
DB 105 VNWGYLAACPSK 117

RESULT 6
AAU75812
ID AAU75812 standard; Protein: 381 AA.

AC AAU75812;

DT 23-APR-2002 (first entry)

XX Wheat L endoxylanase inhibitor, TAXI I; variant #1.

XX wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
KM immobilised enzyme; enzyme; dough syringing; cereal product; beer; plant;
KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KM food; animal feed; starch separation; maize processing; malting;
KM plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX Triticum aestivum.

OS key Location/Qualifiers

FT key MISC-difference 95 /Label= Unknown

FT MISC-difference 98 /Label= Unknown

FT MISC-difference 101 /Label= Unknown

FT MISC-difference 110 /Label= Unknown

FT MISC-difference 333 /Label= Unknown

PN WO200198474-A1.

XX 27-DEC-2001.

PF 21-JUN-2001; 2001WO-BE00106.

PR 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.

PR 16-MAR-2001; 2001GB-0006564.

PR 21-MAY-2001; 2001GB-0012328.

XX (LEUV-) LEUVEN RES & DEV.

XX Delcourt J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;

PI Van Campenhout S;

PT WPI; 2002-114579/15.

PS Claim 127; Page 50-51; 127pp: English.

CC The invention relates to separating and/or isolating inhibitors of

CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises

CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of

CC different specificity or by using an affinity chromatographic step with

CC immobilised enzymes and/or antibodies against inhibitors. Also

CC included are an isolated nucleic acid molecule encoding an inhibitor

CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,

CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,

CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the

CC nucleic acid, an expression system transformed with the nucleic

CC acid, a host organism transformed with the nucleic acid, the inhibitory

CC proteins encoded by the nucleic acids and modulators of the proteins.

CC A recombinant protein, glycoprotein or polypeptide or microorganisms,

CC plant or plant materials transformed with the nucleic acid are useful

CC for the formation of an endoxylanase-inhibitor complex, screening

CC endoxylanases that are totally, less or not inhibited by the inhibitors,

CC reducing syringing in refrigerated dough compositions, affecting the

CC relative affinity and/or relative hydrolysis specificity and/or relative

CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an

CC endoxylanase/inhibitor complex, improving the malting of cereals such as

CC barley, sorghum and wheat and/or the production of beer, improving the

CC production and/or quality of baked or extruded cereal products such as

CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,

CC biscuits, pasta and noodles, animal feed stuff, improving the production

CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten

CC starch separation and production, improving maize processing,

CC plant disease resistance and nutraceutical and/or pharmaceutical

CC applications, improving paper and pulp technologies. The present

CC sequence is a variant of wheat TAXI-I (T. aestivum L endoxylanase

CC inhibitor).

CC Note: Variant amino acids are highlighted in the specification but

CC no wild-type sequence is shown for comparison.

XX SQ Sequence 381 AA:

Query Match 100.0%; Score 65; DB 23; Length 381;

Best Local Similarity 100.0%; Pred. No. 0.00098;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNWGYLAACPSK 13

DB 116 VNWGYLAACPSK 128

RESULT 7

AAU75813

ID AAU75813 standard; Protein: 381 AA.
 AC AAU75813;
 XX
 XX 23-APR-2002 (first entry)
 DT
 DE Wheat L endoxylanase inhibitor, TAXI I, variant #2.
 XX
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KM xylanolytic enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KM immobilised enzyme; sponge dough; Chorleywood bread; biscuits; pasta;
 KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KM noodle; animal feed; starch separation; maize processing; malting;
 KM plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 OS Triticum aestivum.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95
 FT /Label= Unknown
 FT Misc-difference 98
 FT /Label= Unknown
 FT Misc-difference 101
 FT /Label= Unknown
 FT Misc-difference 110
 FT /Label= Unknown
 FT Misc-difference 333
 FT /Label= Unknown
 XX
 XX WO200198474-A1.
 PM
 XX
 XX 27-DEC-2001.
 PD
 XX
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEUVEN RES & DEV.
 PA
 XX Delcourt J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 XX
 DR WPI; 2002-114579/15.
 XX
 XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes -
 XX
 PS Claim 127; Page 51; 127pp; English.
 XX
 XX The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the

CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of wheat TAXI-I (T. aestivum L endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 XX
 SQ Sequence 381 AA:
 XX
 XX
 XX Query Match 100.0%; Score 65; DB 23; Length 381;
 XX Best Local Similarity 100.0%; Pred. No. 0.00098;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VNVGVLAACAPSK 13
 DB 116 VNVGVLAACAPSK 128
 XX
 XX
 XX RESULT 8
 XX AAU75814
 XX ID AAU75814 standard; Protein: 381 AA.
 XX
 XX
 XX AAU75814;
 AC
 XX
 XX 23-APR-2002 (first entry)
 DT
 XX
 XX Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.
 DE
 XX
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KM immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KM noodle; animal feed; starch separation; maize processing; malting;
 KM plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KW
 XX
 XX Triticum aestivum.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95
 FT /Label= Unknown
 FT Misc-difference 98
 FT /Label= Unknown
 FT Misc-difference 101
 FT /Label= Unknown
 FT Misc-difference 110
 FT /Label= Unknown
 FT Misc-difference 145
 FT /Label= Asp, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 183
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 232
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 275
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 282
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 333
 FT /Label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 XX
 XX WO200198474-A1.
 PN
 XX 27-DEC-2001.
 XX

PF 21-JUN-2001; 2001WO-BE00106.
 XX
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 XX (LEUV-) LEUVEN RES & DEV.
 PA
 PI Delcour J, Debysse W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Campenhout S;
 DR MPI; 2002-114579/15.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 51; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a microheterogenic variant of wheat TAXI-I (T. aestivum L
 CC endoxylanase inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 XX Sequence 381 AA:
 Query Match 100.0%; Score 65; DB 23; Length 381;
 Best Local Similarity 100.0%; Pred. No. 0.00098;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VNVGVLAACAPSK 13
 DB 116 VNVGVLAACAPSK 128
 RESULT 9
 AAU75819
 ID AAU75819 standard; Protein; 287 AA.
 XX
 AC AAU75819;

XX
 DT 23-APR-2002 (first entry)
 XX
 DE wheat L endoxylanase inhibitor, partial sequence TAXI-III.
 XX
 KW wheat; TAXI-III; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Triticum aestivum cultivar Soissons.
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 XX
 PI Delcour J, Debysse W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Campenhout S;
 DR MPI; 2002-114579/15.
 XX
 DR N-PSDB; ABR13676.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprising using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 58; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-III (T. aestivum L endoxylanase
 CC inhibitor).

```

XX SQ Sequence 287 AA:
Query Match 76.9%; Score 50; DB 23; Length 287;
Best Local Similarity 76.9%; Pred. No. 0.43;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNWGVLAACAPSK 13
   ||:|||||
Db 107 VNVRAVAAACAPSK 119

RESULT 10
AAU75818
ID AAU75818 standard; Protein; 277 AA.
XX AAU75818;
AC
XX
XX
XX 23-APR-2002 (first entry)
DT
XX
XX L endoxylanase inhibitor, ATXI-II, partial sequence ATXI-II.01.
DE
XX
XX ATXI-II; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor; ds;
KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
KW noodle; animal feed; starch separation; maize processing; malting;
KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
XX
XX Aegilops tauschii variety Acc2220051.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 4 /note= "Encoded by CTC"
FT
XX
XX WO200198474-A1.
XX
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-BE00106.
XX
XX 22-JUN-2000; 2000GB-0015296.
XX 25-JAN-2001; 2001GB-0002018.
XX 26-JAN-2001; 2001GB-0002194.
XX 16-MAR-2001; 2001GB-0006564.
XX 21-MAY-2001; 2001GB-0012328.
XX
XX (LEUV-) LEUVEN RES & DEV.
XX
XX Delcour J, Debysse W, Gebruers K, Goesaert H, Fierens K, Robben J;
PI Van Campenhout S;
XX
XX WPI: 2002-114579/15.
XX N-PSDB: ABK13675.
XX
XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
PT beta-glucanolytic enzymes comprises using endoxylanases during
PT screening for inhibition activity or affinity chromatography with
PT immobilised enzymes
XX
XX Claim 127; Page 58; 127pp; English.
XX
XX The invention relates to separating and/or isolating inhibitors of
XX cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
XX screening the inhibition activity by using two or more enzymes during the
XX separation and/or isolation steps that allow to distinguish inhibitors of
XX different specificity or by using an affinity chromatographic step with
XX immobilised enzymes and/or antibodies against inhibitors. Also
XX included are an isolated nucleic acid molecule encoding an inhibitor
XX which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
XX alpha-L-arabino-furanosidase and/or other cellulose, xylan,
XX arabinoxylan or beta-galactan degrading enzymes, a vector comprising the
XX nucleic acid, an expression system transformed with the nucleic

```

```

CC acid, a host organism transformed with the nucleic acid, the inhibitory
CC proteins encoded by the nucleic acids and modulators of the proteins.
CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
CC plant or plant materials transformed with the nucleic acid are useful
CC for the formation of an endoxylanase-inhibitor complex, screening
CC endoxylanases that are totally, less or not inhibited by the inhibitors,
CC reducing syruping in refrigerated dough compositions, affecting the
CC relative affinity and/or relative hydrolysis specificity and/or relative
CC hydrolysis rate versus water-extractable and/or water-unextractable
CC arabinoxylans of endoxylanases such as by the formation of an
CC endoxylanase/inhibitor complex, improving the malting of cereals such as
CC barley, sorghum and wheat and/or the production of beer, improving the
CC production and/or quality of baked or extruded cereal products such as
CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
CC biscuits, pasta and noodles, animal feed stuff, improving the production
CC of starch derived syrups, sorbitol, xylitol, wheat gluten
CC starch separation and production, improving maize processing,
CC plant disease resistance and nutraceutical and/or pharmaceutical
CC applications, improving paper and pulp technologies. The present
CC sequence represents a partial Aegilops tauschii ATXI-II (Aegilops
CC tauschii L endoxylanase inhibitor).
XX
XX SQ Sequence 277 AA:
Query Match 66.2%; Score 43; DB 23; Length 277;
Best Local Similarity 66.2%; Pred. No. 8.3;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNWGVLAACAPSK 13
   ||:|||||
Db 105 VNVRIILAACAPSK 117

RESULT 11
AAW81605
ID AAW81605 standard; Peptide; 25 AA.
XX
XX AAW81605;
AC
XX
XX 27-JAN-1999 (first entry)
DT
XX
XX Synthetic rabies nucleoprotein homologue peptide fragment.
XX
XX 9D: vaccine; cytotoxic T cell response; viral infection; rabies;
KW peptide-fatty acid conjugate; influenza; oncogene; liposome; APC;
KW human immune deficiency virus; antigen-presenting cell; antibody;
KW protection.
XX
XX Synthetic.
XX
XX US5837249-A.
XX
XX 17-NOV-1998.
XX
XX 20-OCT-1993; 93US-0139609.
XX
XX 15-APR-1992; 92US-0868946.
XX 19-APR-1985; 85US-0725087.
XX 08-MAY-1987; 87US-0047443.
XX 12-APR-1991; 91US-0685459.
XX 20-OCT-1993; 93US-0139609.
XX
XX (WIST-) WISTAR INST.
XX
XX Dietzschold B, Heber-Katz E;
PI
XX
XX WPI: 1999-023378/02.
XX
XX Inducing cytotoxic T cell response against virus using peptide-fatty
PT acid conjugate - Formulated in liposomes with an adjuvant,
PT specifically for protecting against herpes simplex or rabies viruses
XX
XX Example 11; Column 14; 21pp; English.

```

XX This sequence is a synthetic peptide fragment which is a homologue of the
 CC rabies virus N protein. This sequence is used in a method of inducing
 CC a cytotoxic T cell response against viral infection in a mammal by
 CC administering a peptide-fatty acid conjugate. Such conjugates are used
 CC particularly to vaccinate against HSV, rabies and also other viruses such
 CC as influenza, human immune deficiency virus and oncogenic viruses. When
 CC the liposomes fuse to an antigen-presenting cell (APC), the conjugate
 CC remains bound to the surface of the APC membrane and is not degraded
 CC inside the cell, which generates a T cell response without any antibody
 CC response, avoiding the risk of immune enhancement (in which antibodies
 CC increase viral infectivity). Such compounds can provide long-lasting
 CC protection from only a single injection.

CC Sequence 25 AA;

Query Match 63.1%; Score 41; DB 20; Length 25;
 Best Local Similarity 72.7%; Pred. No. 1.4;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGLAACAP 11
 : | : |||||
 DB 13 LNATVIAACAP 23

RESULT 12

AAP80639
 ID AAP80689 standard; peptide; 25 AA.

AC AAP80689;

DT 07-NOV-1990 (first entry)

DE Sequence corresp. to a peptide fragment derived from rabies
 DE virus nucleoprotein.

KW Immunodominant T-cell antigen; rabies vaccines;
 KW peptide-fatty acid conjugate.

OS Rabies virus.

PN EP290246-A.

PD 09-NOV-1988.

PF 05-MAY-1988; 88EP-0304045.

PR 08-MAY-1987; 87US-0047443.

PA (WIST-) WISTAR INST ANATOMY.

PI Heber-Katz E, Dietzschold B;

DR WPI; 1988-316484/45.

PT Peptide-fatty acid conjugates;
 PT used in vaccines for generating an immunogenic T cell response
 PT protective against a virus disease state e.g. rabies

PS Claim 6; Page 22; 36pp; English.

CC Peptide fragment is an immunodominant T cell antigenic determinant. It
 CC is part of a new peptide-fatty acid conjugate. Also claimed is a vaccine
 CC for generating an immunogenic T cell response protective against a virus
 CC disease state comprising:
 CC (a) a peptide-fatty acid complex or a synthetic replica;
 CC (b) a liposome compsn. comprising a mixt. of phosphatidyl choline,
 CC cholesterol and lysophosphatidyl cholines, (proportions by wt. pref.
 CC 16:2:1); and
 CC (c) an adjuvant.

CC Vaccines achieve significant protection for a protracted period of time
 CC against a large dose of virus by a single injection. A peptide having
 CC sequence homology to the 25 AA peptide chain of rabies nucleoprotein,

CC having a spacer of Gly-Gly-Lys added to N-terminus and Cys added to the
 CC C-terminus and having the formula in AAP80689 was prepd. One month after
 CC a single immunisation (10ug peptide/animal), mice were challenged with a
 CC lethal dose of rabies virus (2 LD₅₀) dose of strain CVS-24 grown in BHK
 CC cells) and the animals examined for the next 21 days for paralysis and
 CC death; 87% of vaccine injected animals survived.

CC Sequence 25 AA;

Query Match 60.0%; Score 39; DB 9; Length 25;
 Best Local Similarity 63.6%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGLAACAP 11
 : | : |||||
 DB 13 LNATVIAACAP 23

RESULT 13

AAP91478
 ID AAP91478 standard; protein; 25 AA.

AC AAP91478;

DT 24-JUN-1990 (first entry)

DE Peptide N-V12b.

KW Rabies related virus; peptide N-V12b; vaccine.

OS Rabies related virus.

PN WO8900861-A.

PD 09-FEB-1989.

PF 29-JUL-1988; 88WO-US02529.

PR 30-JUL-1987; 87US-0079639.

PA (WIST-) THE WISTAR INSTITUTE.

PI Dietzschold B, Koprowski H;

DR WPI; 1989-061041/08.

PT New rabies virus polypeptide(s) - useful for vaccine prodn.

PS Claim 13; page 25; 31pp; English.

CC Peptide N-V10b corresponds to amino acids 313-337 of rabies related virus
 CC (including rabies virus and related viruses such as Mokola and European
 CC bat virus) nucleoprotein. It is useful for the prodn of vaccines against
 CC homologous and heterologous rabies virus strains. They can be produced
 CC without growing the virus, and do not cause the adverse side effects of
 CC inactivated virus vaccines, eg demyelinating allergic encephalitis and
 CC systemic reactions. The peptide is opt. administered in liposomes
 CC (coupled to 15-21C fatty acids beforehand), followed by at least one
 CC booster injection, esp. of inactivated virus or a mixt. of virus
 CC glycoprotein and nucleoprotein.

CC Sequence 25 AA;

Query Match 60.0%; Score 39; DB 10; Length 25;
 Best Local Similarity 63.6%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGLAACAP 11
 : | : |||||
 DB 13 LNATVIAACAP 23

RESULT 14

AAW81603
ID AAW81603 standard; peptide; 25 AA.
XX
AC AAW81603;
XX
DT 27-JAN-1999 (first entry)
XX
DE Synthetic peptide fragment N-V12b.
XX
KW gD; vaccine; cytotoxic T cell response; viral infection; rabies;
KW peptide-fatty acid conjugate; influenza; oncogene; liposome; APC;
KW human immune deficiency virus; antigen-presenting cell; antibody;
KW protection.
XX
OS Synthetic.
XX
PN US837249-A.
XX
PD 17-NOV-1998.
XX
PF 20-OCT-1993; 93US-0139609.
XX
PR 15-APR-1992; 92US-0868946.
PR 19-APR-1985; 85US-0725087.
PR 08-MAY-1987; 87US-0047443.
PR 12-APR-1991; 91US-0685459.
PR 20-OCT-1993; 93US-0139609.
XX
PA (WIST-) WISTAR INST.
XX
PI Dietzschold B, Heber-Katz E;
XX
DR WPI; 1999-023378/02.
XX
PT Inducing cytotoxic T cell response against virus using peptide-fatty
PT acid conjugate - formulated in liposomes with an adjuvant,
PT specifically for protecting against herpes simplex or rabies viruses
XX
PS Claim 21; Column 4; 21pp; English.
XX
CC This sequence is a synthetic peptide fragment, N-V12b, which is a
CC homologue of the rabies virus N protein. This sequence is used in a
CC method of inducing a cytotoxic T cell response against viral infection in
CC a mammal by administering a peptide-fatty acid conjugate. Such conjugates
CC are used particularly to vaccinate against HSV, rabies and also other
CC viruses such as influenza, human immune deficiency virus and oncogenic
CC viruses. When the liposomes fuse to an antigen-presenting cell (APC), the
CC conjugate remains bound to the surface of the APC membrane and is not
CC degraded inside the cell, which generates a T cell response without any
CC antibody response, avoiding the risk of immune enhancement (in which
CC antibodies increase viral infectivity). Such compounds can provide
CC long-lasting protection from only a single injection.
XX
SQ Sequence 25 AA;
XX
Query Match 60.0%; Score 39; DB 20; Length 25;
Best Local Similarity 63.6%; Pred. No. 3.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 VNWGVLAACAP 11
! !!!!!
DB 13 LNATVIAACAP 23
XX
RESULT 15
AAG42355
ID AAG42355 standard; Protein; 341 AA.
XX
AC AAG42355;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52814.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 15-JUL-1999; 99US-0144334.
PR 15-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145092.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 23-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158362.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 60.0%; Score 39; DB 21; Length 341;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 2 NNGVLAACPSK 13
DB 3 SIVGTALCSPAK 14

Search completed: May 20, 2003, 16:14:43
Job time : 15.1429 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 5.28571 seconds
(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-18
Perfect score: 65
Sequence: 1 VNVGVLAACAPSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	63.1	169	2 C75427	hypothetical prote
2	40	61.5	228	2 F75390	hypothetical prote
3	40	61.5	349	2 F95263	selenide, water di
4	40	61.5	493	2 S11148	ama protein - Str
5	40	61.5	659	2 A95221	hypothetical prote
6	40	61.5	659	2 A98085	hypothetical prote
7	39	60.0	257	2 AC2406	2-hydroxyhepta-2,4
8	39	60.0	288	2 T24066	exonuclease III ho
9	39	60.0	290	2 T09363	hypothetical prote
10	39	60.0	409	2 S12018	endoglucanase B -
11	39	60.0	437	2 J02015	nucleoprotein - ra
12	39	60.0	437	2 J02016	nucleoprotein - ra
13	39	60.0	450	1 VHNVRY	nucleoprotein - ra
14	39	60.0	450	1 VHNVNSB	nucleoprotein - ra
15	39	60.0	450	1 VHNVAV	nucleoprotein - ra
16	38	58.5	131	2 A87461	hypothetical prote
17	38	58.5	208	2 F70734	hypothetical prote
18	38	58.5	384	2 AH0241	ceramide glucosylt
19	38	58.5	835	2 C86444	hypothetical prote
20	37	56.9	147	2 S77485	ribosomal protein
21	37	56.9	263	2 D71281	probable pyroline
22	37	56.9	334	2 D87575	hypothetical prote
23	37	56.9	409	2 F70633	probable lipk prot
24	37	56.9	431	2 T50573	probable integral
25	37	56.9	450	2 A58460	nucleoprotein N -
26	37	56.9	459	2 D70914	hypothetical prote
27	37	56.9	1117	2 A38227	RNA-splicing regul
28	36	55.4	155	2 E5174	hypothetical 16.7
29	36	55.4	247	2 T35602	probable transposa

30	36	55.4	359	2 A13259	hypothetical prote
31	36	55.4	385	2 G71474	probable cell divi
32	36	55.4	396	2 A13554	xylose transport s
33	36	55.4	407	2 H81736	cell division prot
34	36	55.4	471	2 D83388	probable MFS trans
35	36	55.4	559	2 H84859	hypothetical prote
36	36	55.4	738	2 B96744	unknown protein [i
37	36	55.4	621	2 C84700	hypothetical prote
38	35.5	54.6	621	2 H87652	feruloyl-coa synth
39	35	53.8	128	2 T28433	selenophosphate sy
40	35	53.8	179	2 S63610	probable prophage
41	35	53.8	179	2 JC6047	regB protein - Ser
42	35	53.8	182	2 E83220	conserved hypothet
43	35	53.8	210	2 C87256	hypothetical prote
44	35	53.8	217	2 S62795	probable lipoprote
45	35	53.8	278	2 G81139	hypothetical prote

ALIGNMENTS

RESULT 1
C75427
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: C75427
R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75427
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-169 <WHI>
A:Cross-references: GB:AE001967; GB:AE00513; NID:g6458915; PIDN:AAF10759.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1180
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1180

Query Match 63.1% Score 41; DB 2; Length 169;
Best Local Similarity 70.0%; Pred. No. 6.9;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGVLAACAPS 12
DB 13 VGLASCAPA 22

RESULT 2
F75390
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75390
R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75390
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-228 <WHI>
A:Cross-references: GB:AE001992; GB:AE00513; NID:g6459233; PIDN:AAF11048.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1474

A:Map position: 1

Query Match
Best Local Similarity 61.5%; Score 40; DB 2; Length 228;
Matches 7; Conservative 77.8%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 GVLAACAP 11
11:11:111
DB 14 GVLAACAP 22

RESULT 3
F95263
seleninfe, water dikinase (EC 2.7.9.3) seld [imported] - Shnorhizobium meliloti (strain 1
N:Alternate names: selenophosphate synthetase
C:Species: Shnorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F95263
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; ; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Shnorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: F95263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK64672.1; PID:g14523070; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hymn, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebbaut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Shnorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: seld
A:Genome: plasmid
C:Keywords: transferase

Query Match
Best Local Similarity 61.5%; Score 40; DB 2; Length 349;
Matches 7; Conservative 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GVLAACAPSK 13
1:11111111
DB 307 GLTVACAPSK 316

RESULT 4
S11148
amla protein - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C:Accession: S11148
R:Alloing, G.; Trombe, M.C.; Claverys, J.P.
Mol. Microbiol. 4, 633-644, 1990
A:Title: The aml locus of the Gram-positive bacterium Streptococcus pneumoniae is similar
A:Reference number: S11148; MUID:90279506; PMID:2352474
A:Accession: S11148
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <ALL>
A:Cross-references: EMBL:X17337

Query Match
Best Local Similarity 61.5%; Score 40; DB 2; Length 493;
Matches 8; Conservative 80.0%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GVLAACAPSK 13

DB 18 GVLAACSSSK 27
111111:11

RESULT 5
A95221
hypothetical protein SP1891 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: A95221
R:Rettelli, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11465916
A:Accession: A95221
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75962.1; PID:g14973394; GSPDB:GN00164; TIGR
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1891

Query Match
Best Local Similarity 61.5%; Score 40; DB 2; Length 659;
Matches 8; Conservative 80.0%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GVLAACAPSK 13
111111:11
DB 18 GVLAACSSSK 27

RESULT 6
A98085
hypothetical protein amla [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: A98085
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A98085
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00510.1; PID:g15459383; GSPDB:GN00174
C:Genetics:
A:Gene: amla

Query Match
Best Local Similarity 61.5%; Score 40; DB 2; Length 659;
Matches 8; Conservative 80.0%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GVLAACAPSK 13
111111:11
DB 18 GVLAACSSSK 27

RESULT 7
AC2406
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase [imported] - Nostoc sp. (strain PCC 712
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AC2406
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AC2406
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <KUR>
 A:Cross-References: GB:BA000019; PIDN:BAW76502.1; PID:g17133940; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr4803
 C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 60.0%; Score 39; DB 2; Length 257;
 Best Local Similarity 77.8%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 VLAACAPSK 13
 :|||
 Db 51 ILAPCAPSK 59

RESULT 8
 T24066
 exonuclease III homolog - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T24066; T42391
 R:Kershaw, J.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19837
 A:Accession: T24066
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-288 <MIL>
 A:Cross-References: EMBL:281108; PIDN:CA03235.1; GSPDB:GN00019; CESP:R09B3.1
 A:Experimental source: clone R09B3
 R:Tremblay, S.; Masson, J.Y.; Ramotar, D.
 submitted to the EMBL Data Library, November 1997
 A:Description: The exonuclease III family of DNA repair enzymes has a new homolog in *Ce*
 A:Reference number: Z22150
 A:Accession: T42391
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-288 <PRE>
 A:Cross-References: EMBL:AF034258; PIDN:AAC82328.1
 C:Genetics:
 A:Gene: R09B3.1
 A:Map position: 1
 A:Introns: 83/3; 192/3; 264/3
 C:Superfamily: exodeoxyribonuclease III

Query Match 60.0%; Score 39; DB 2; Length 288;
 Best Local Similarity 63.6%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 VGVLAACAPSK 13
 ||:|:|
 Db 102 VGLSKCAPSK 112

RESULT 9
 T09363
 hypothetical protein F23K16.100 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C:Accession: T09363
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16652
 A:Accession: T09363
 A:Molecule type: DNA

A:Residues: 1-290 <BEV>
 A:Cross-References: EMBL:AL078620; GSPDB:GN00062; ATSP:F23K16.100
 A:Experimental source: cultivar Columbia; BAC clone F23K16
 C:Genetics:
 A:Gene: ATSP:F23K16.100
 A:Map position: 4
 A:Introns: 205/3; 253/3

Query Match 60.0%; Score 39; DB 2; Length 290;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 NVGVLAACAPSK 13
 :|||
 Db 3 SIGVTALCSPAK 14

RESULT 10
 S12018
 endoglucanase B - *Ruminococcus albus*
 C:Species: *Ruminococcus albus*
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
 C:Accession: S12018
 R:Poole, D.M.; Hazlewood, G.P.; Laurie, J.I.; Barker, P.J.; Gilbert, H.J.
 Mol. Gen. Genet. 223, 217-223, 1990
 A:Title: Nucleotide sequence of the *Ruminococcus albus* SY3 endoglucanase genes *celA* a
 A:Reference number: S12017; MUID:91066833; PMID:2250649
 A:Accession: S12018
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <POO>
 A:Cross-References: GB:X54932; NID:g45965; PIDN:CAA38693.1; PID:g45966
 C:Superfamily: *Ruminococcus albus* cellulase

Query Match 60.0%; Score 39; DB 2; Length 409;
 Best Local Similarity 53.8%; Pred. No. 33;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGVLAACAPSK 13
 :|||:|
 Db 15 MSVGVMAACGCGSK 27

RESULT 11
 J02015
 nucleoprotein - rabies virus (isolate T2)
 C:Species: rabies virus
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: J02015
 R:Nadin-Davis, S.A.; Casey, G.A.; Wandeler, A.
 J. Gen. Virol. 74, 829-837, 1993
 A:Title: Identification of regional variants of the rabies virus within the Canadian
 A:Reference number: J02012; MUID:93260396; PMID:8492088
 A:Accession: J02015
 A:Molecule type: genomic RNA
 A:Residues: 1-437 <NAD>
 C:Genetics:
 A:Gene: N
 C:Superfamily: rabies virus nucleoprotein
 C:Keywords: nucleoprotein

Query Match 60.0%; Score 39; DB 2; Length 437;
 Best Local Similarity 63.6%; Pred. No. 35;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGVLAACAP 11
 :|||
 Db 320 LNAVLAACAP 330

RESULT 12
 J02016
 nucleoprotein - rabies virus (isolate T3)

C:Species: rabies virus
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: JQ2016
R:Nadin-Davis, S.A.; Casey, G.A.; Mandeler, A.
J. Gen. Virol. 74, 829-837, 1993
A:Title: Identification of regional variants of the rabies virus within the Canadian pro
A:Reference number: JQ2012; MUID:93260396; PMID:8492088
A:Accession: JQ2016
A:Molecule type: genomic RNA
A:Residues: 1-437 <NAD>
C:Genetics:
A:Gene: N
C:Superfamily: rabies virus nucleoprotein
C:Keywords: nucleoprotein

Query Match 60.0%; Score 39; DB 2; Length 437;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNWGLAACAP 11
: | | | | |
Db 320 LNATVIAACAP 330

RESULT 13

VHVNRY
nucleoprotein - rabies virus
C:Species: rabies virus
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D26275; A24887
R:Tordo, N.; Poch, O.; Ermline, A.; Keith, G.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 83, 3914-3918, 1986
A:Title: Walking along the rabies genome: is the large G-L intergenic region a remnant
A:Reference number: A94100; MUID:86233343; PMID:3459163
A:Accession: D26275
A:Molecule type: genomic RNA
A:Residues: 1-450 <TOR>
A:Cross-references: GB:M13215; GB:M21634; NID:9333585; PIDN:AAA47215.1; PID:9333586
C:Genetics:
A:Gene: N
C:Superfamily: rabies virus nucleoprotein
C:Keywords: nucleoprotein

Query Match 60.0%; Score 39; DB 1; Length 450;
Best Local Similarity 63.6%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNWGLAACAP 11
: | | | | |
Db 325 LNATVIAACAP 335

RESULT 14

VHVNRS
nucleoprotein - rabies virus (strain SAD B19)
C:Species: rabies virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: A34746
R:Conzelmann, K.K.; Cox, J.H.; Schneider, L.G.; Thiel, H.J.
Virology 175, 485-499, 1990
A:Title: Molecular cloning and complete nucleotide sequence of the attenuated rabies vir
A:Reference number: A34746; MUID:90223994; PMID:2139267
A:Accession: A34746
A:Molecule type: genomic RNA
A:Residues: 1-450 <CON>
A:Cross-references: GB:M31046; NID:9333556; PIDN:AAA47199.1; PID:9333557
C:Genetics:
A:Gene: N
C:Superfamily: rabies virus nucleoprotein
C:Keywords: nucleoprotein

Query Match 60.0%; Score 39; DB 1; Length 450;
Best Local Similarity 63.6%; Pred. No. 36;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 VNWGLAACAP 11
: | | | | |
Db 325 LNATVIAACAP 335

RESULT 15

VHVNAY
nucleoprotein - rabies virus (strain AVO1)
C:Species: rabies virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: S07813
R:Poch, O.; Tordo, N.; Keith, G.
Biochimie 70, 1019-1029, 1988
A:Title: Sequence of the 3386 3' nucleotides of the genome of the AVO1 strain rabies
A:Reference number: S07813; MUID:89150295; PMID:3147658
A:Accession: S07813
A:Molecule type: genomic RNA
A:Residues: 1-450 <POC>
A:Cross-references: EMBL:X13357; NID:961809; PIDN:CAA31733.1; PID:961810
C:Genetics:
A:Gene: N
C:Superfamily: rabies virus nucleoprotein
C:Keywords: nucleoprotein

Query Match 60.0%; Score 39; DB 1; Length 450;
Best Local Similarity 63.6%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNWGLAACAP 11
: | | | | |
Db 325 LNATVIAACAP 335

Search completed: May 20, 2003, 16:37:06
Job time: 7.28571 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 3.5 Seconds

(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-18

Sequence: 1 VNVGVLAACAPSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	61.5	659	1 AMIA_STRPN	P18791 streptococ
2	39	60.0	409	1 GUNB_RUMAL	P23661 rumalnococu
3	39	60.0	450	1 NCAP_RABVA	P15197 rabies viru
4	39	60.0	450	1 NCAP_RABVP	P08314 rabies viru
5	39	60.0	450	1 NCAP_RABVP	P06025 rabies viru
6	39	60.0	450	1 NCAP_RABVP	P16285 rabies viru
7	39	60.0	450	1 NCAP_RABVP	P09110 rabies viru
8	38	58.5	150	1 PAL_PASMU	P05186 pasteurella
9	38	58.5	208	1 YN06_MYCTU	P51886 pasteurella
10	37	56.9	147	1 RL15_SYNY3	P73303 mycobacteri
11	37	56.9	263	1 PROC_TREPA	P27771 treponema p
12	37	56.9	459	1 YE43_MYCTU	P71694 mycobacteri
13	37	56.9	1117	1 CYT4_NEUCR	P47950 neuropept
14	36	55.4	154	1 SARA_STRCO	P12690 streptomyc
15	36	55.4	155	1 YIEI_ECOLI	P31468 escherichia
16	36	55.4	372	1 CD14_RABIT	P02880 cryotolagus
17	35	53.8	217	1 YF90_MYCPN	P05037 mycoplasma
18	35	53.8	517	1 MURF_MYCLE	P06956 mycobacteri
19	35	53.8	663	1 GNP1_YEAST	P48813 saccharomyc
20	35	53.8	887	1 OBP_HSEVB	P28947 equine harp
21	35	53.8	942	1 HEX_ADEG1	P42671 avian adeno
22	35	53.8	955	1 MML3_MYCLE	P06081 mycobacteri
23	35	53.8	2556	1 NML1_HUMAN	P46531 homo sapien
24	34	52.3	55	1 SARA_STRCO	P13106 streptococ
25	34	52.3	173	1 CRAB_RANCA	P04694 guillardia
26	34	52.3	204	1 RK3_GUTTH	P40131 salmonella
27	34	52.3	219	1 FLGA_SALTY	P19641 escherichia
28	34	52.3	323	1 ISPB_ECOLI	P04266 lactulur p
29	34	52.3	346	1 OPRP_ICTPU	P00737 homo sapien
30	34	52.3	347	1 HPT1_HUMAN	P50417 ateles geof
31	34	52.3	347	1 HPT1_HUMAN	P50417 ateles geof
32	34	52.3	348	1 HPT1_HUMAN	P50417 ateles geof
33	34	52.3	362	1 KLF1_HUMAN	P13351 homo sapien

34	34	52.3	375	1 CD14_HUMAN	P08571 homo sapien
35	34	52.3	403	1 MHPT_ECOLI	P77589 escherichia
36	34	52.3	406	1 HPT2_HUMAN	P00738 homo sapien
37	34	52.3	407	1 MOSC_RHIME	P07609 rhizobium m
38	34	52.3	414	1 FTSW_ECOLI	P16457 escherichia
39	34	52.3	473	1 DCOR_CANAL	P78599 candida alb
40	34	52.3	484	1 SCAT_ECOLI	P39301 escherichia
41	34	52.3	751	1 UAS3_DROME	P56676 drosophila
42	34	50.8	33	1 TXH1_SELHU	P56676 selennocomi
43	33	50.8	158	1 RM19_YEAST	P53875 saccharomyc
44	33	50.8	176	1 KITH_SFVKA	P07605 shope fibro
45	33	50.8	214	1 PCTT_RAT	P53809 rattus norv

ALIGNMENTS

RESULT 1	ID	AMIA_STRPN	STANDARD:	PRT:	659 AA.
AC	P18791	AMIA_STRPN	STANDARD:	PRT:	659 AA.
DT	01-NOV-1990	(Rel. 16, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Oligopeptide-binding protein amia precursor.				
GN	AMIA OR SPI891.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1313;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=R800:				
RX	MEDLINE=90279506; PubMed=2352474;				
RA	Allong G., Trombe M.C., Claverys J.-P.;				
RT	"The ami locus of the Gram-positive bacterium Streptococcus pneumoniae				
RT	is similar to binding protein-dependent transport operons of gram-				
RT	negative bacteria.";				
RL	Mol. Microbiol. 4:633-644(1990).				
RN	(2)				
RP	REVISIONS.				
RA	Claverys J.-P.;				
RL	Submitted (SEP-1992) to the EMBL/Genbank/DBJ databases.				
RN	(3)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TIGR4;				
RX	MEDLINE=21357209; PubMed=11463916;				
RA	Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,				
RA	Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,				
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,				
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,				
RA	Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,				
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,				
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,				
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser J.C.M.;				
RT	"Complete genome sequence of a virulent isolate of Streptococcus				
RT	pneumoniae.";				
RT	Science 293:498-506(2001).				
RN	(4)				
RP	SEQUENCE OF 1-28 FROM N.A.				
RC	STRAIN=R800:				
RX	MEDLINE=90060772; PubMed=2684766;				
RA	Martin B., Allong G., Boucraut C., Claverys J.-P.;				
RT	"The difficulty of cloning Streptococcus pneumoniae mal and ami loci				
RT	in Escherichia coli: toxicity of malx and amia gene products.";				
RT	Gene 80:227-238(1989).				
RN	(5)				
RP	PARTIAL SEQUENCE.				
RX	MEDLINE=89091147; PubMed=3208757;				
RA	Gilson E., Allong G., Schmidt T., Claverys J.-P., Dudler R.,				
RA	Hofnung M.;				
RT	"Evidence for high affinity binding-protein dependent transport				
RT	systems in Gram-positive bacteria and in Mycoplasma.";				

RL EMBL J. 7:3971-3974(1988).
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR OLIGOPEPTIDES, PROBABLY AN OLIGOPEPTIDE BINDING PROTEIN.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 5.
CC -1- CAUTION: THE REVISED SEQUENCE OF AMIA NOW INCLUDES, IN THE
CC C-TERMINAL SECTION, THE SEQUENCE OF AN ORF WHICH WAS PREVIOUSLY
CC KNOWN AS AMIB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17337; CAA35213.1; -
CC EMBL; AE007479; AAK75962.1; -
CC PIR; S11148; S11148.
CC PIR; S11149; S11149.
CC TIGR; SP1891; -
CC InterPro: IPR000914; SBP_bac-5.
CC DR Pfam: PF00496; SBP_bac-5; 1.
CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.
CC KW Peptide transport; Transport; Signal; Membrane; Lipoprotein;
CC Complete proteome.
CC FT SIGNAL 1 22 PROBABLE.
CC FT CHAIN 1 659 OLIGOPEPTIDE-BINDING PROTEIN AMIA.
CC FT LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).
CC FT CONFLICT 54 54 S -> R (IN REF. 1).
CC FT CONFLICT 421 421 A -> R (IN REF. 1).
CC FT CONFLICT 557 557 R -> K (IN REF. 1).
CC FT CONFLICT 562 562 L -> F (IN REF. 1).
CC FT CONFLICT 622 622 S -> L (IN REF. 1).
CC FT CONFLICT 626 626 V -> A (IN REF. 1).
CC SQ SEQUENCE 659 AA; 72465 MW; C2ED06A1DB8A7B81 CRC64;
Query Match 61.5%; Score 40; DB 1; Length 659;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 GVLACAPSK 13
Db 18 GVLACSSK 27
RESULT 2
GUNB_RUMAL STANDARD; PRT; 409 AA.
AC P23661;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endoglucanase B precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
DE (Cellulase) (ECB).
GN cELB.
OS Ruminooccus albus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminooccus.
CC Ruminooccus.
OX NCBI_TaxID=1264;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SY3;
RX MEDLINE=91066833; PubMed=2250649;
RA Poole D.M., Hazlewood G.P., Laurie J.I., Barker P.J., Gilbert H.J.;
RT "Nucleotide sequence of the Ruminooccus albus SY3 endoglucanase
genes celsa and celsb".
KW Mol. Genet. 223:217-223(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic

CC linkages in cellulose.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54932; CAA38693.1; -
CC PIR; S12018; S12018.
CC HSSP; P17901; 1EDG.
CC DR InterPro: IPR001547; GH_5.
CC DR Pfam: PF00150; cellulase; 1.
CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.
CC DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
CC KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT CHAIN 1 409 ENDOLUCANASE B.
CC FT ACT_SITE 212 212 PROTON DONOR (BY SIMILARITY).
CC FT ACT_SITE 332 332 NUCLEOPHILE (BY SIMILARITY).
CC SQ SEQUENCE 409 AA; 45523 MW; 119081DFA3BFD54 CRC64;
Query Match 60.0%; Score 39; DB 1; Length 409;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 VNVGLACAPSK 13
Db 15 MSVGVMAVSCGSK 27
RESULT 3
NCAP_RABVA STANDARD; PRT; 450 AA.
AC P15197;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Nucleocapsid protein (Nucleoprotein).
GN N.
OS Rabies virus (strain AVOI).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11293;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89150295; PubMed=3147698;
RA Poch O., Tordo N., Keith G.;
RT "Sequence of the 3386 3' nucleotides of the genome of the AVOI strain
RT rabies virus: structural similarities in the protein regions involved
RT in transcription".
RT Biochimie 70:1019-1029(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13357; CAA31733.1; -
CC PIR; S07813; VHVNAV.
CC DR InterPro: IPR000448; Rhabd_nucleocap.
CC DR Pfam: PF00945; Rhabd_nucleocap; 1.
CC DR ProDom: PD002087; Rhabd_nucleocap; 1.
CC KW Nucleocapsid.
CC SQ SEQUENCE 450 AA; 50733 MW; 00493D2751305A76 CRC64;

DR EMBL: M31046; AAA47199.1; -
 DR PIR: A34746; VHVNSB.
 DR InterPro: IPR000448; Rhabd_nucleocap.
 DR Pfam: PF00945; Rhabd_nucleocap; 1.
 DR ProDom: PD002087; Rhabd_nucleocap; 1.
 KW Nucleocapsid.
 SQ SEQUENCE 450 AA; 50603 MW; 783BF3E01E7BE325 CRC64;

Query Match 60.0%; Score 39; DB 1; Length 450;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGLAACAP 11
 : | : | | | | |
 Db 325 LNATVIAACAP 335

RESULT 7

NCAP_RABVU STANDARD; PRT; 450 AA.

AC QC9110;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DE 01-FEB-1995 (Rel. 31, Last annotation update)

DE Nucleocapsid protein (Nucleoprotein).

GN N.

OS Rabies virus (strain Ontario skunk).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Lyssavirus.

OX NCBI_TaxID=39005;

RN [..]

RP SEQUENCE FROM N.A.

RX MEDLINE=93260396; PubMed=8492088;

RA Nadin-Davis S.A.; Casey G.A.; Wandeler A.;

RT "Identification of regional variants of the rabies virus within the

RT Canadian province of Ontario";

RL J. Gen. Virol. 74:829-837(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL: L20671; AAA47220.1; -

DR InterPro: IPR000448; Rhabd_nucleocap.

DR Pfam: PF00945; Rhabd_nucleocap; 1.

DR ProDom: PD002087; Rhabd_nucleocap; 1.

KW Nucleocapsid.

SQ SEQUENCE 450 AA; 50611 MW; D0FACAA9A7773C69 CRC64;

Query Match 60.0%; Score 39; DB 1; Length 450;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VNVGLAACAP 11
 : | : | | | | |
 Db 325 LNATVIAACAP 335

RESULT 8

PAL_PASMU

ID PAL_PASMU STANDARD; PRT; 150 AA.

AC Q51886;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Outer membrane protein P6 precursor (OMP P6) (P6-like) (Peptidoglycan-

DE associated lipoprotein).

DE

OS PASTEURELLA multocida.

OS

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Pasteurella.

OX NCBI_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-T16;

RX MEDLINE=95172751; PubMed=7868272;

RA Kasten R.W.; Hansen L.M.; Hinojoza J.; Bieber D.; Ruehl W.W.;

RT Hirsch D.C.;

RT "Pasteurella multocida produces a protein with homology to the P6

RT outer membrane protein of Haemophilus influenzae";

RL Infect. Immun. 63:989-993(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J.; Zhang Q.; Li L.L.; Paustian M.L.; Whitam T.S.; Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID

CC ANCHOR (By similarity).

CC -1- SIMILARITY: TO OTHER PAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL: U16849; AAA79373.1; -

DR EMBL: AE006136; AAK03050.1; -

DR InterPro: IPR001145; Bac-OmpA.

DR Pfam: PF00691; OmpA; 1.

DR PRINTS: PR01021; OMPADOMAIN.

DR ProDom: PD000930; Bac-OmpA; 1.

DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.

DR PROSITE: PS01068; OMPA; 1.

KW Outer membrane; Signal; Lipoprotein; Complete proteome.

FT SIGNAL 1 19

FT CHAIN 20 150

FT LIPID 20 20

FT DOMAIN 82 126

FT OMPA-LIKE.

SQ SEQUENCE 150 AA; 16213 MW; 170E7B13D2E9ED6C CRC64;

Query Match 58.5%; Score 38; DB 1; Length 150;
 Best Local Similarity 66.7%; Pred. No. 6.6;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 NNGVLAACAPSK 13
 : | | | | | | |
 Db 13 SVAVLAACGSSK 24

RESULT 9
 YN06_MYCTU STANDARD; PRT; 208 AA.

AC Q50659;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein RV2306C precursor.

DE RV2306C OR MT2363 OR MTCV339.03.

GN Mycobacterium tuberculosis.

OS Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37Rv;

RX MEDLINE=98293987; PubMed=9634230;

RA Cole S.T.; Brosch R.; Parkhill J.; Garnier T.; Churcher C.; Harris D.;

RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Razaizadeh M.A., Rogers J.,
RA Rutter S., Seeger J., Skellon S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.,
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickley E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A.L., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 277163; CAB00992.1;
DR EMBL: AE007078; AAK46649.1;
DR TIGR: MT2363;
DR TubercuList: RV2306c;
DR Hypothetical protein: Signal: Complete proteome.
KW SIGNAL
FT CHAIN 17 208
FT SEQUENCE 208 AA; 22521 MW; DE913A5D2C34DA2F CRC64;
SQ
Query Match 58.5%; Score 38; DB 1; Length 208;
Best Local Similarity 54.5%; Pred. No. 8.8;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 3 VGVLAACAPSK 13
DB 131 LGVTPACAPAE 141
RESULT 10
RL15_SYNY3
ID RL15_SYNY3 STANDARD; PRT; 147 AA.
AC P73303;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE 50S ribosomal protein L15.
GN RPL0 OR RPL15 OR SL1813.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose A., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -I- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA

CC (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90905; BAA17332.1;
DR InterPro: IPR001196; Ribosomal_L15.
DR Pfam: PF00256; L15; 1.
DR Pfam: PF01305; Ribosomal_L15; 1.
DR TIGRPFAM: TRG01071; rpl0_bact; 1.
DR PROSITE: PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 147 AA; 15195 MW; 18DC0B2F69B1AE09 CRC64;
Query Match 56.9%; Score 37; DB 1; Length 147;
Best Local Similarity 66.7%; Pred. No. 9.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VNVGVLAACAPS 12
DB 78 VNVGKLAGLAPN 89
RESULT 11
PROC TREPA
ID PROC_TREPA STANDARD; PRT; 263 AA.
AC P27771; O83775;
DT 01-AUG-1992 (rel. 23, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Pyroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PROC OR TP0797.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90264287; PubMed=2188947;
RA Cheradimi F.C., Hobbs M.M., Stamm L.V., Bassford P.J. Jr.;
RT "Complementation of an Escherichia coli proc mutation by a gene cloned
RT from Treponema pallidum.";
RL J. Bacteriol. 172:2996-3002(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -I- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = L-pyrroline-5-
CC carboxylate + NAD(P)H.
CC -I- PATHWAY: Proline biosynthesis; third (last) step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR HMBL: M73825; AAA2478.1; -
 DR HMBL: AE001250; AAC65760.1; -
 DR PIR: S27734; S27734.
 DR TIGR: TP0797; -
 DR InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR.1.
 DR PROSITE: PS00521; P5CR.1.
 DR OXidoreductase; Proline biosynthesis; NADP; Complete proteome.
 FT CONFLICT 1 65
 FT FT
 FT CONFLICT 77 93
 FT CONFLICT 256 260
 FT CONFLICT 263 AA; 27645 MW; ED4AD7C54BAF9D61 CRC64;
 SO SEQUENCE

Query Match
 Best Local Similarity 56.9%; Score 37; DB 1; Length 263;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GVLAACAPS 12
 Db 169 GVINGCAPA 177

RESULT 12
 YE43_MYCTU STANDARD; PRT; 459 AA.
 ID YE43_MYCTU
 AC P71694; 006833;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV1425.
 GN RV1425 OR MT1468 OR MTCY21B4.43 OR MTCY493.29C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98295987; Pubmed=9634230;
 RX STRAIN=H37RV;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 295844; CAB09245.1; -
 DR EMBL: AE007017; AAK45733.1; ALT_INIT.
 DR TIGR: MT1468; -
 DR TubercuList: Rv1425; -
 DR InterPro: IPR004235; UPF0089.
 DR Pfam: PF03007; UPF0089.1.
 DR Hypothetical protein; Complete proteome.
 SO SEQUENCE 459 AA; 50062 MW; F436D6B58562263 CRC64;

Query Match
 Best Local Similarity 56.9%; Score 37; DB 1; Length 459;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNVGVLAACAPS 12
 Db 21 MHVGAALICDPS 32

RESULT 13
 CYT4_NEUCR STANDARD; PRT; 1117 AA.
 ID CYT4_NEUCR
 AC P47950;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Mitochondrial protein cyt-4.
 GN CYT-4.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A;
 RX MEDLINE=92179246; Pubmed=1311848;
 RA Turc B., Dobinson K.F., Serizawa N., Lambowitz A.M.;
 RT "A protein required for RNA processing and splicing in Neurospora
 RT mitochondria is related to gene products involved in cell cycle
 RT protein phosphatase functions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1676-1680(1992).
 CC -1- FUNCTION: REQUIRED FOR RNA 5' AND 3' END PROCESSING AND SPLICING.
 CC MAY ACT ON THE RNA PROCESSING ENZYMES DIRECTLY, OR IT MAY ACT ON
 CC OTHER REGULATORY MOLECULES, WHICH INFLUENCE THE ACTIVITY OR
 CC SYNTHESIS OF THESE ENZYMES.
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEASE II (RNB) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M80735; AAA33608.1; -
 DR PIR: A38227; A38227.
 DR InterPro: IPR001900; Ribonuclease_II.
 DR Pfam: PF00773; RNB.1.
 DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
 KM Mitochondrion; mRNA processing.
 SO SEQUENCE 1117 AA; 125103 MW; 93D1832D9128B20B CRC64;

Query Match
 Best Local Similarity 56.9%; Score 37; DB 1; Length 1117;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NCVLAACAPS 12
1:|||||1|
Db 204 NCVLAACAPS 214

RESULT 14
SAPA_STRCO STANDARD; PRT; 154 AA.
AC P12690; Q9RJS7;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Spore-associated protein A precursor.
GN SAPA OR SC00409 OR SCF51.08C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieger H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieger T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajadream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Witzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
RN [2]
RP SEQUENCE OF 1-116 FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=86169521; PubMed=2450872;
RA Guirao J., Santamaria R., Schauer A., Losick R.;
RT "Promoter determining the timing and spatial localization of
transcription of a cloned Streptomyces coelicolor gene encoding a
spore-associated polypeptide.";
RT J. Bacteriol. 170:1895-1901(1988).
CC -|- SURCELLULAR LOCATION: SPORE SURFACE.
CC -|- DEVELOPMENTAL STAGE: TRANSCRIPTION OF THIS PROTEIN IS INDUCED AT
THE TIME OF APPEARANCE OF AERIAL MYCELIIUM BEFORE THE BEGINNING OF
SPORE FORMATION.
CC -|- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 79
ONWARD DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A1132707; CAB59706.1; -;
DR EMBL; M20145; AAA26812.1; ALT_FRAME.
KW Sporulation; Signal; Complete proteome.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 154 SPORE-ASSOCIATED PROTEIN A.
SQ SEQUENCE 154 AA; 15641 MW; B0B72EAF54361547 CRC64;

Query Match 55.4%; Score 36; DB 1; Length 154;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCVLAACAPS 13
1:|||||1|

Db 136 NCGALAKAPOK 147

RESULT 15
YIEL_ECOLI STANDARD; PRT; 155 AA.
ID YIEL_ECOLI
AC P31468;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yiel.
GN YIEL OR B3716.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication.";
RL Genomics 16:551-561(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L10328; AAA62067.1; -;
DR EMBL; AE000448; AAC76739.1; -;
DR Ecogene; EGI1726; yiel.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 155 AA; 16683 MW; 779FDA7162CF501D CRC64;

Query Match 55.4%; Score 36; DB 1; Length 155;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GVALACAPS 13
1:|||||1|
Db 63 GVALACAPS 72

Search completed: May 20, 2003, 16:35:46
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 16.0714 Seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-18
Perfect score: 65
Sequence: 1 VNVGVLAACAPSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	63.1	169	16 Q9RV50	Q9RV50 deinoococcus
2	41	63.1	419	5 Q9NA38	Q9NA38 caenorhabdi
3	41	63.1	420	5 Q9NA61	Q9NA61 caenorhabdi
4	41	63.1	450	12 Q80Y66	Q80Y66 rabies viru
5	40	61.5	228	16 Q9RUB6	Q9RUB6 deinoococcus
6	40	61.5	349	16 Q931D0	Q931D0 rhizobium m
7	40	61.5	637	5 Q9GZ30	Q9GZ30 trypanosoma
8	39	60.0	257	16 Q8YMX1	Q8YMX1 anabaena sp
9	39	60.0	288	5 Q45711	Q45711 caenorhabdi
10	39	60.0	290	10 Q9SVB1	Q9SVB1 arabidopsis
11	39	60.0	341	10 Q94A56	Q94A56 rabies viru
12	39	60.0	445	12 Q99106	Q99106 rabies viru
13	39	60.0	450	12 Q85412	Q85412 rabies viru
14	39	60.0	450	12 Q55611	Q55611 rabies viru
15	39	60.0	450	12 Q9W9S4	Q9W9S4 rabies viru
16	39	60.0	450	12 Q39473	Q39473 rabies viru

17	39	60.0	450	12 Q86407	Q86407 rabies viru
18	39	60.0	450	12 Q86408	Q86408 rabies viru
19	39	60.0	450	12 Q86409	Q86409 rabies viru
20	39	60.0	450	12 Q86410	Q86410 rabies viru
21	39	60.0	450	12 Q86413	Q86413 rabies viru
22	39	60.0	450	12 Q86414	Q86414 rabies viru
23	39	60.0	450	12 Q86415	Q86415 rabies viru
24	39	60.0	450	12 Q86416	Q86416 rabies viru
25	39	60.0	450	12 Q86417	Q86417 rabies viru
26	39	60.0	450	12 Q86418	Q86418 rabies viru
27	39	60.0	450	12 Q86420	Q86420 rabies viru
28	39	60.0	450	12 Q86421	Q86421 rabies viru
29	39	60.0	450	12 Q86422	Q86422 rabies viru
30	39	60.0	450	12 Q86423	Q86423 rabies viru
31	39	60.0	450	12 Q86424	Q86424 rabies viru
32	39	60.0	450	12 Q86425	Q86425 rabies viru
33	39	60.0	450	12 Q86426	Q86426 rabies viru
34	39	60.0	450	12 Q86427	Q86427 rabies viru
35	39	60.0	450	12 Q86429	Q86429 rabies viru
36	39	60.0	450	12 Q86430	Q86430 rabies viru
37	39	60.0	450	12 Q86431	Q86431 rabies viru
38	39	60.0	450	12 Q86432	Q86432 rabies viru
39	39	60.0	450	12 Q86433	Q86433 rabies viru
40	39	60.0	450	12 Q86434	Q86434 rabies viru
41	39	60.0	450	12 Q86435	Q86435 rabies viru
42	39	60.0	450	12 Q86436	Q86436 rabies viru
43	39	60.0	450	12 Q86437	Q86437 rabies viru
44	39	60.0	450	12 Q86438	Q86438 rabies viru
45	39	60.0	450	12 Q86441	Q86441 rabies viru

ALIGNMENTS

RESULT 1

Q9RV50 ID Q9RV50 PRELIMINARY; PRT; 169 AA.
AC Q9RV50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DR1180.
GN DR1180.
OS Deinoococcus radiodurans.
OC Bacteria; Thermus/Deinoococcus group; Deinoococi; Deinoococcales;
OC Deinoococaceae; Deinoococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "genome sequence of the radioresistant bacterium Deinoococcus
RT radiodurans RI.";
DR Science 286:1571-1577(1999).
DR EMBL; AF001967; AAF10759.1; -
DR TIGR; DR1180; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 169 AA; 17313 MW; 22D92DC79C5A0C5 CRC64;

Query Match Best Local Similarity 70.0%; Score 41; DB 16; Length 169;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGVLAACAPS 12
11:11:11:
Db 13 VGLIASCAPA 22

RESULT 2

Q9N38 PRELIMINARY; PRT; 419 AA.
 AC Q9N38; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 GN Y3F8A.9 protein.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RT [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AL132862; CAB70227.1; -;
 DR InterPro; IPR003341; DUF139.
 DR Pfam; PF02363; DUF139; 11.
 SQ SEQUENCE 419 AA; 45397 MW; F62317E9E7B408DF CRC64;

Query Match

Best Local Similarity 63.1%; Score 41; DB 5; Length 419;
 Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGTLACAP 11

Db 194 VNLGISAQCAP 204

RESULT 3

Q9N61 PRELIMINARY; PRT; 420 AA.
 AC Q9N61; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 GN Y3F8A.8 protein.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RT [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AL132862; CAB60533.1; -;
 DR InterPro; IPR003341; DUF139.
 DR Pfam; PF02363; DUF139; 11.
 SQ SEQUENCE 420 AA; 45412 MW; 51BAB4ED5EA9E538 CRC64;

Query Match 63.1%; Score 41; DB 5; Length 420;
 Best Local Similarity 63.6%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGTLACAP 11
 Db 195 VNLGISAQCAP 205

RESULT 4

Q80Y66 PRELIMINARY; PRT; 450 AA.
 AC Q80Y66; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 GN Nucleoprotein.
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jacobs J.A.; Theron J.; Nel L.H.;
 RT "In vitro RNA- and protein-binding characteristics of the
 nucleoprotein of a South African viverrid rabies virus isolate."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF467949; AAL76066.1; -;
 SQ SEQUENCE 450 AA; 50606 MW; E834525A67FC7D74 CRC64;

Query Match

Best Local Similarity 63.1%; Score 41; DB 12; Length 450;
 Pred. No. 23;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGTLACAP 11

Db 325 LMAVIAACAP 335

RESULT 5

Q9RUB6 PRELIMINARY; PRT; 228 AA.
 AC Q9RUB6; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 GN Hypothetical protein DR1474.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX SPRAIN-RI;
 RC MEDLINE=20036896; PubMed=10567266;
 RA White O.; Eisen J.A.; Heidelberg J.F.; Hickey E.K.; Peterson J.D.;
 RA Dodson R.J.; Haft D.H.; Gwinn M.L.; Nelson W.C.; Richardson D.L.;
 RA Moffat K.S.; Qin H.; Jiang L.; Pamphile W.; Crosby M.; Shen M.;
 RA Vamathevan J.J.; Lam P.; McDonald L.; Utterback T.; Zalewski C.;
 RA Makarova K.S.; Aravind L.; Daly M.J.; Minton K.W.; Fleischmann R.D.;
 RA Ketchum K.A.; Nelson K.E.; Salzberg S.; Smith H.O.; Venter J.C.;
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans RI."
 RL Science 286:1571-1577(1999).
 DR EMBL; AEO01992; AAF11048.1; -;
 DR TIGR; DR1474; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 228 AA; 24659 MW; D26BA136F57A4087 CRC64;

Query Match

Best Local Similarity 61.5%; Score 40; DB 16; Length 228;
 Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGVIAACAP 11

Db 11:11:11

Db 14 VGLLSCAP 22

RESULT 6

Q931D0 PRELIMINARY; PRT; 349 AA.
 ID 0931D0
 AC 0931D0;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Selenide, water dikinase (selenophosphate synthetase), Seld (EC 2.7.9.3).
 GN Seld OR RA0014 OR SMA0028.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymba (megaplasmid 1).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huzar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kallman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSymba megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL: AE007196; AAK64672.1; -;
 DR InterPro: IPR000728; AIRS.related.
 DR InterPro: IPR004536; Seld.
 DR Pfam: PF02769; AIRS.C.1.
 DR TIGRfams: TIGR00476; Seld; 1.
 DR Kinase; Transferase; Plasmid; Complete proteome.
 SQ SEQUENCE 349 AA; 35833 MW; 1196A5D9D6F87CF3 CRC64;

Query Match

Best Local Similarity 61.5%; Score 40; DB 16; Length 349;
 Pred. No. 27;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GVLACAPSK 13
 Db 307 GLVACAPSK 316

RESULT 7

Q9GZ30 PRELIMINARY; PRT; 637 AA.
 ID 09GZ30
 AC 09GZ30;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical 72.2 kDa protein.
 OS Trypanosoma cruzi.
 CC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 NCBI_TaxID=5693;
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Andersson B.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF242860; AAF98148.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 637 AA; 72214 MW; 6470347EF695D33F CRC64;

Query Match

Best Local Similarity 61.5%; Score 40; DB 5; Length 637;
 Pred. No. 49;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 NCVLAACAPS 12
 I :|||||:

Db 473 NTNVAACAPT 483

RESULT 8

O8YMX1 PRELIMINARY; PRT; 257 AA.
 ID 08YMX1
 AC 08YMX1;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase.
 GN ALR4803.
 OS Anabaena sp. (strain PCC 7120).
 CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CC NCBI_TaxID=103690;
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneo T., Nakamura Y., Wolk C.P.,
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003597; BAB76502.1; -;
 DR InterPro: IPR002529; FAA_hydrolase.
 DR Pfam: PF01557; FAA_hydrolase; 1.
 KW Isomerase; Complete proteome.
 SQ SEQUENCE 257 AA; 28446 MW; 7B9F7470042323A5 CRC64;

Query Match

Best Local Similarity 60.0%; Score 39; DB 16; Length 257;
 Pred. No. 31;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VLAACAPSK 13
 Db 51 ILAPCAPSK 59

RESULT 9

O45711 PRELIMINARY; PRT; 288 AA.
 ID 045711
 AC 045711;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE R09B3.1 protein.
 GN R09B3.1.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kershaw J.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., Momurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailton N., Smith A., Sonhammer E., Straden R., Sulston J.,
 RA Telford-Mleg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RL Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tremblay S., Masson J.-Y., Ramotar D.;
 RT "The exonuclease III family of DNA repair enzymes has a new homolog in
 CEEXIII from *Caenorhabditis elegans*."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z81108; CAB03235.1; -
 DR EMBL: AF034258; AAC82328.1; -
 DR HSSP: P27695; 1E9N.
 DR InterPro: IPR004442; Exonase_III.
 DR InterPro: IPR004808; ExoIII_xth.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR TIGRFAMs: TIGR00195; exonase_III; 1.
 DR TIGRFAMs: TIGR00633; xth; 1.
 KW Exonuclease.
 SQ SEQUENCE 288 AA; 32898 MW; 104C97DE2B7F9661 CRC64;

Query Match 60.0%; Score 39; DB 5; Length 288;
 Best Local Similarity 63.6%; Pred. No. 35;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VGVLAACAPSK 13
 : : : : : : :
 DB 102 VGLSKCAPMK 112

RESULT 10

Q9SVBI PRELIMINARY; PRT; 290 AA.
 AC Q9SVBI;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Hypothetical 33.4 kDa protein.
 GN F23K16.100 OR AT4G39470.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eurosids II: Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL078620; CAB44682.1; -
 DR EMBL: AL161595; CAB80610.1; -
 DR InterPro: IPR001440; TPR.
 KW Hypothetical protein.
 SQ SEQUENCE 290 AA; 33444 MW; A76771155D9E72A5 CRC64;

Query Match 60.0%; Score 39; DB 10; Length 290;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 NVGVLAACAPSK 13
 : : : : : : :
 DB 3 SIGVTALCAPSK 14

RESULT 11

Q94A56 PRELIMINARY; PRT; 341 AA.
 AC Q94A56;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE AT4G39470/F23K16.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eurosids II: Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY050356; MAK91373.1; -
 DR InterPro: IPR004440; TPR.
 DR SMART: SM00028; TPR; 1.
 SWART: SM00028; TPR; 1.
 SQ SEQUENCE 341 AA; 39155 MW; 9708D7021C06B5E1 CRC64;

Query Match 60.0%; Score 39; DB 10; Length 341;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 NVGVLAACAPSK 13
 : : : : : : :
 DB 3 SIGVTALCAPSK 14

RESULT 12

Q99106 PRELIMINARY; PRT; 445 AA.
 AC Q99106;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Nucleocapsid protein (Nucleoprotein) (Fragment).
 OS Rabies virus.
 OC Viruses: ssRNA negative-strand viruses: Mononegavirales;
 OC Rhabdoviridae: Lyssavirus.
 NCBI_TaxID=11292;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306453; PubMed=1840709;
 RA Sumner J.W., Pekada M., Shaddock J.H., Esposito J.J., Bellini W.;
 RT "Protection of mice with vaccinia virus recombinants that express the
 rabies nucleoprotein."
 RL Virology 185:703-710(1991).
 DR EMBL: M61047; AAA47222.1; -
 DR InterPro: IPR000448; Rhabd_nucleocap.
 DR Pfam: PF00945; Rhabd_nucleocap; 1.
 DR Prodom: PD002087; Rhabd_nucleocap; 1.
 KW Nucleocapsid.

FT NON_TER
 SQ SEQUENCE 445 AA; 50183 MW; 76734579132E3F00 CRC64;

Query Match 60.0%; Score 39; DB 12; Length 445;
 Best Local Similarity 63.6%; Pred. No. 53;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVGVLAACAPSK 11
 : : : : : : :
 DB 1 VGVLAACAPSK 11

Db 320 LNATVIAACAP 330

RESULT 13

085412 PRELIMINARY; PRT; 450 AA.
 AC Q85412;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE NUCLEOPROTEIN.
 GN N.
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RC-HL;
 RC MEDLINE=94353646; PubMed=8073640;
 RA Goto H., Minamoto N., Ito H., Sugiyama M., Kinjo T., Mannen K.,
 RA Mitune K., Kawato A.;
 RT "Nucleotide sequence of the nucleoprotein gene of the RC-HL strain of
 RT rabies virus, a seed strain used for animal vaccine production in
 RT Japan."
 RT Virus genes 8:91-97(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RC-HL;
 RA Minamoto N.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RC-HL;
 RA Minamoto N.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D16331; BAA03838.1; -
 DR EMBL: AB009663; BAA24083.1; -
 DR InterPro: IPR000448; Rhabd_nucleocap.
 DR Pfam: PF00945; Rhabd_nucleocap; 1.
 DR ProDom: PD002087; Rhabd_nucleocap; 1.
 KW Nucleoprotein.
 SQ SEQUENCE 450 AA; 50658 MW; A73F8EB623D8A5A9 CRC64;

Query Match 60.0%; Score 39; DB 12; Length 450;
 Best Local Similarity 63.6%; Pred. No. 53;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGVLAACAP 11
 : | : |||||
 Db 325 LNATVIAACAP 335

RESULT 14

055611 PRELIMINARY; PRT; 450 AA.
 AC 055611;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE NUCLEOPROTEIN.
 GN N.
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NISHIGAHARA;
 RA Minamoto N.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-NISHIGAHARA;

RA Ito N., Kakemizu M., Ito K.A., Yamamoto A., Yoshida Y., Sugiyama M.,

RA Minamoto N.;

RT "A comparison of complete genome sequences of the attenuated RC-HL

RT strain of rabies virus used for production of animal vaccine in Japan,

RT and the parental Nishigahara strain."

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB010494; BAA24618.1; -

DR EMBL: AB044824; BAA96802.1; -

DR InterPro: IPR000448; Rhabd_nucleocap.

DR Pfam: PF00945; Rhabd_nucleocap; 1.

DR ProDom: PD002087; Rhabd_nucleocap; 1.

KW Nucleoprotein.

SQ SEQUENCE 450 AA; 50720 MW; 7DAFB96AA7055A2C CRC64;

QY 1 VNVGVLAACAP 11
 : | : |||||
 Db 325 LNATVIAACAP 335

RESULT 15

09W9S4 PRELIMINARY; PRT; 450 AA.
 AC 09W9S4;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE NUCLEOPROTEIN.
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=86G1FRA; AND 9223FRA;
 RA Bouthy H., Kissi B., Kulonen K., Tordo N., Audry L., Stohr K.;
 RT "Evolution of the nucleoprotein gene and host range of old world
 RT rabies virus."
 RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RL EMBL: U43434; AAD10458.1; -
 DR EMBL: U43433; AAD10457.1; -
 DR InterPro: IPR000448; Rhabd_nucleocap.
 DR Pfam: PF00945; Rhabd_nucleocap; 1.
 DR ProDom: PD002087; Rhabd_nucleocap; 1.
 KW Nucleocapsid.
 SQ SEQUENCE 450 AA; 50662 MW; DCAC028FFBDAF689 CRC64;

Query Match 60.0%; Score 39; DB 12; Length 450;
 Best Local Similarity 63.6%; Pred. No. 53;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGVLAACAP 11
 : | : |||||
 Db 325 LNATVIAACAP 335

Search completed: May 20, 2003, 16:43:20
 Job time : 18.0714 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 : Search time 6 Seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-18
Perfect score: 65
Sequence: 1 VNWGVLAACAPSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5A.COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5B.COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	63.1	25	2	US-08-139-609-5
2	40	61.5	642	2	US-08-245-511-48
3	40	61.5	642	2	US-08-600-993A-48
4	39	60.0	25	2	US-08-139-609-3
5	38	58.5	26	2	US-08-139-609-12
6	36	55.4	1422	4	US-08-469-260A-83
7	36	55.4	2864	4	US-08-469-260A-394
8	35	53.8	52	4	US-09-461-697-95
9	35	53.8	485	4	US-09-516-914-1
10	35	53.8	569	4	US-08-686-968C-227
11	35	53.8	749	4	US-09-562-737-99
12	35	53.8	942	4	US-09-171-461-12
13	35	53.8	2556	1	US-08-185-432-17
14	35	53.8	2556	1	US-08-083-590A-20
15	35	53.8	2556	3	US-08-532-384-20
16	35	53.8	2556	3	US-08-899-232-2
17	34	52.3	334	2	US-08-484-397A-8
18	34	52.3	348	1	US-08-366-953A-45
19	34	52.3	348	2	US-08-484-397A-2
20	34	52.3	348	2	US-08-484-397A-3
21	34	52.3	348	2	US-08-484-397A-4
22	34	52.3	348	2	US-08-484-397A-5
23	34	52.3	348	2	US-08-484-397A-6
24	34	52.3	348	2	US-08-484-397A-7
25	34	52.3	348	2	US-08-484-397A-27
26	34	52.3	348	2	US-08-484-397A-38
27	34	52.3	359	4	US-09-134-001C-3728

28	34	52.3	361	4	US-08-874-569B-21	Sequence 21, Appl
29	34	52.3	375	1	US-08-205-719-2	Sequence 2, Appl
30	34	52.3	375	3	US-08-746-883-5	Sequence 5, Appl
31	34	52.3	643	2	US-08-245-511-47	Sequence 47, Appl
32	34	52.3	643	2	US-08-600-993A-47	Sequence 47, Appl
33	34	52.3	643	2	US-08-637-759B-441	Sequence 441, App
34	33	50.8	114	2	US-08-871-355A-441	Sequence 441, App
35	33	50.8	114	4	US-09-201-945-441	Sequence 441, App
36	33	50.8	326	3	US-09-154-874-9	Sequence 9, Appl
37	33	50.8	562	3	US-08-427-097-30	Sequence 30, Appl
38	33	50.8	562	2	US-08-878-957-30	Sequence 30, Appl
39	33	50.8	564	1	US-08-427-097-2	Sequence 2, Appl
40	33	50.8	564	1	US-08-427-097-14	Sequence 14, Appl
41	33	50.8	564	1	US-08-427-097-16	Sequence 16, Appl
42	33	50.8	564	1	US-08-427-097-20	Sequence 20, Appl
43	33	50.8	564	1	US-08-427-097-28	Sequence 28, Appl
44	33	50.8	564	2	US-08-878-957-2	Sequence 2, Appl
45	33	50.8	564	2	US-08-878-957-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-139-609-5
Sequence 5, Application US/08139609
Patent No. 5837249
GENERAL INFORMATION:
APPLICANT: Heber-Katz, Ellen
TITLE OF INVENTION: Method for Generating an Immunogenic T
TITLE OF INVENTION: Cell Response Protective Against a Virus
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457
City: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/139,609
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,946
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/685,459
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/047,443
FILING DATE: 08-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/725,087
FILING DATE: 19-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-139-609-5

Query Match 63.1%; Score 41; DB 2; Length 25;
Best Local Similarity 72.7%; Pred. No. 0.53;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGLAACAP 11
: | | | | | | | |
Db 13 LNAVTLAACAP 23

RESULT 2
US-08-245-511-48

Sequence 48, Application US/08245511
Patent No. 5928900

GENERAL INFORMATION:

APPLICANT: Masure, H Robert

APPLICANT: Pearce, Barbara J

APPLICANT: Tuomanen, Elaine

TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/245,511

FILING DATE: 18-MAY-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/116,541

FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-069 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 642 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

IMMEDIATE SOURCE:

CLONE: am1a

FEATURE: OTHER INFORMATION: NOTE: the reference contains a

OTHER INFORMATION: sequence error; the correct sequence shown below is obtained

PUBLICATION INFORMATION:

AUTHORS: Allouing, et al.

JOURNAL: Mol. Microbiol.

VOLUME: 4

PAGES: 633-644

DATE: 1990

US-08-245-511-48

Query Match 61.5%; Score 40; DB 2; Length 642;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GVLAAAPSK 13
: | | | | | | | |
Db 1 GVLAAACSSK 10

RESULT 3
US-08-600-993A-48

Sequence 48, Application US/08600993A
Patent No. 5981229

GENERAL INFORMATION:

APPLICANT: Masure, H Robert

APPLICANT: Pearce, Barbara J

APPLICANT: Tuomanen, Elaine

TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/600,993A

FILING DATE: 1-MAR-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/245,511

FILING DATE: 18-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/116,541

FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-069 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 642 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

IMMEDIATE SOURCE:

CLONE: am1a

FEATURE: OTHER INFORMATION: the reference contains a sequence error; the

OTHER INFORMATION: correct sequence shown below is obtained from GENBANK

PUBLICATION INFORMATION:

AUTHORS: Allouing, et al.

JOURNAL: Mol. Microbiol.

VOLUME: 4

PAGES: 633-644

DATE: 1990

US-08-600-993A-48

US-08-600-993A-48

Query Match

61.5%; Score 40; DB 2; Length 642;

Best Local Similarity 80.0%; Pred. No. 27;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GVLAACPSK 13

Db 1 GVLAACSSK 10

RESULT 4

US-08-139-609-3

; Sequence 3, Application US/08139609

; Patent No. 5837249

; GENERAL INFORMATION:

; APPLICANT: Heber-Katz, Ellen

; APPLICANT: Dietzschold, Bernhard

; TITLE OF INVENTION: Method for Generating an Immunogenic T

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr., P.O. Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/139,609

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/868,946

; FILING DATE: 15-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/685,459

; FILING DATE: 12-APR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/047,443

; FILING DATE: 08-MAY-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 06/725,087

; FILING DATE: 19-APR-1985

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: WSTIDUSA

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ. ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; US-08-139-609-3

US-08-139-609-3

Query Match

60.0%; Score 39; DB 2; Length 25;

Best Local Similarity 63.6%; Pred. No. 1.2;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNVGLAACAP 11

Db 13 LNAVIVAACAP 23

RESULT 5

US-08-139-609-12

; Sequence 12, Application US/08139609

; Patent No. 5837249

; GENERAL INFORMATION:

; APPLICANT: Heber-Katz, Ellen

; APPLICANT: Dietzschold, Bernhard

; TITLE OF INVENTION: Method for Generating an Immunogenic T

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr., P.O. Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/139,609

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/868,946

; FILING DATE: 15-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/685,459

; FILING DATE: 12-APR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/047,443

; FILING DATE: 08-MAY-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 06/725,087

; FILING DATE: 19-APR-1985

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: WSTIDUSA

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ. ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 26 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; US-08-139-609-12

US-08-139-609-12

Query Match

58.5%; Score 38; DB 2; Length 26;

Best Local Similarity 70.0%; Pred. No. 1.9;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 NVGLAACAP 11

Db 15 NATVIVAACAP 24

RESULT 6

US-08-469-260A-83

; Sequence 83, Application US/08469260A

; Patent No. 6451578

; GENERAL INFORMATION:

; APPLICANT: JOHN N. SIMONS

; APPLICANT: TAMU J. PILOT-MATIAS

; APPLICANT: GEORGE J. DAWSON

; APPLICANT: GEORGE G. SCHLAUDER

; APPLICANT: SURESH M. DESAI

; APPLICANT: THOMAS P. LEARY

APPLICANT: ANTHONY SCOTT MUEHROFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,260A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: POREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-260A-83

Query Match 55.4%; Score 36; DB 4; Length 1422;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGVLAACA 10
|11111111|
DB 783 VGVLSACA 790

RESULT 7
US-08-469-260A-394
Sequence 394, Application US/08469260A
Patent No. 6451578
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUEHROFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK

STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,260A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: POREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 394:
SEQUENCE CHARACTERISTICS:
LENGTH: 2864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-260A-394

Query Match 55.4%; Score 36; DB 4; Length 2864;
Best Local Similarity 87.5%; Pred. No. 7.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGVLAACA 10
|11111111|
DB 1787 VGVLSACA 1794

RESULT 8
US-09-461-697-95
Sequence 95, Application US/09461697
Patent No. 6277974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Puranam, Kasturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
TITLE OF INVENTION: CELL DEATH
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 95
LENGTH: 52
TYPE: PRT
ORGANISM: Homo sapiens
US-09-461-697-95

Query Match 53.8%; Score 35; DB 4; Length 52;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 VLAACAPSK 13
|11111111|
DB 4 VLAACSPSR 12


```
RESULT 9
US-09-516-914-1
; Sequence 1, Application US/09516914
; Patent No. 6333401
; GENERAL INFORMATION:
; APPLICANT: Breinly, Sabine
; APPLICANT: Fuchs, Georg
; TITLE OF INVENTION: Phenol-induced Proteins of Thauera aromatica
; FILE REFERENCE: BC1006 US NA
; CURRENT APPLICATION NUMBER: US/09/516,914
; EARLIER FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 60/122,952
; EARLIER FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Thauera aromatica
US-09-516-914-1

Query Match
Best Local Similarity 53.8%; Score 35; DB 4; Length 485;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VGVLAACAPS 12
; :|||||
Db 211 LAAMAACAPS 220

RESULT 10
US-08-686-968C-227
; Sequence 227, Application US/08686968C
; Patent No. 6221361
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; FILE REFERENCE: 39119-H/JML
; CURRENT APPLICATION NUMBER: US/08/686,968C
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 227
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Swinepox virus
US-08-686-968C-227

Query Match
Best Local Similarity 53.8%; Score 35; DB 4; Length 569;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NCVLAAC 9
; :|||||
Db 149 NIGVTAAC 156

RESULT 11
US-09-562-737-99
; Sequence 99, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
```

```
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-99

Query Match
Best Local Similarity 53.8%; Score 35; DB 4; Length 749;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLAACAPSK 13
; :||:|:|
Db 268 ILASCSPSR 276

RESULT 12
US-09-171-461-12
; Sequence 12, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Mathew
; APPLICANT: Chioocca, Susana
; APPLICANT: Kurzbaue, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 942
; TYPE: PRT
; ORGANISM: CELO virus
; FEATURE:
; OTHER INFORMATION: Position: 18289..21117 /gene: L3 /product: L3 hexon
US-09-171-461-12

Query Match
Best Local Similarity 53.8%; Score 35; DB 4; Length 942;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VNVGLAACAPS 12
; |||:|:|
Db 278 VNSCTMARCRPN 289

RESULT 13
US-08-185-432-17
; Sequence 17, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DETEX PROTEINS, NUCLEIC ACIDS, AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-06-185-432-17

Query Match
Best local Similarity 53.8%; Score 35; DB 1; Length 2556;
Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 VLACAPS 12
11111111
DB 829 VLACAPS 836

RESULT 14
US-06-083-590A-20
Sequence 20, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-083-590A-20

Query Match
Best local Similarity 53.8%; Score 35; DB 1; Length 2556;
Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 VLACAPS 12
11111111
DB 829 VLACAPS 836

RESULT 15
US-08-532-384-20
Sequence 20, Application US/08532384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-532-384-20

Query Match
Best local Similarity 53.8%; Score 35; DB 3; Length 2556;
Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 VLACAPS 12
11111111
DB 829 VLACAPS 836

Search completed: May 20, 2003, 16:44:44
Job time : 7 secs

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 : Search time 6.42857 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-18
Perfect score: 65
Sequence: 1 VNVGVLAACAPSK 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.dep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.dep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.dep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.dep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.dep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.dep:*
7: /cgn2_6/ptodata/1/pubpaa/PCT08_PUBCOMB.dep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.dep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.dep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.dep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.dep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.dep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.dep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.dep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	58.5	35	10	US-09-864-761-37459
2	38	58.5	63	9	US-10-092-154-737
3	38	58.5	63	10	US-09-764-847-737
4	36	55.4	1422	8	US-08-424-550B-83
5	36	55.4	2862	10	US-09-742-659-5
6	36	55.4	2864	8	US-08-424-550B-394
7	36	55.4	2865	10	US-09-742-659-2
8	35	53.8	2865	10	US-09-922-261-95
9	35	53.8	485	10	US-09-870-162A-1
10	35	53.8	749	9	US-10-211-962-99
11	35	53.8	942	9	US-09-970-711-12
12	35	53.8	2444	10	US-09-944-849-2
13	34	52.3	23	9	US-10-097-065-597
14	34	52.3	135	9	US-10-097-065-596
15	34	52.3	143	9	US-09-925-299-1335
16	34	52.3	143	10	US-09-925-299-1335
17	34	52.3	143	10	US-10-072-349-90
18	34	52.3	210	9	US-09-764-855-90
19	34	52.3	210	10	US-09-764-855-90

20	34	52.3	258	10	US-09-864-761-37821	Sequence 37821, A
21	34	52.3	323	10	US-09-815-242-10321	Sequence 10321, A
22	34	52.3	323	10	US-09-815-242-13734	Sequence 13734, A
23	34	52.3	330	9	US-09-738-626-3767	Sequence 3767, Ap
24	34	52.3	346	9	US-09-977-577-8	Sequence 8, Appli
25	34	52.3	347	9	US-09-977-577-1	Sequence 1, Appli
26	34	52.3	347	10	US-09-977-577-3	Sequence 21, Appli
27	34	52.3	361	10	US-09-955-518-21	Sequence 2, Appli
28	34	52.3	406	9	US-09-977-577-2	Sequence 453, App
29	34	52.3	414	10	US-09-741-669-453	Sequence 10030, A
30	34	52.3	414	10	US-09-815-242-10030	Sequence 14077, A
31	34	52.3	414	10	US-09-815-242-14077	Sequence 485, App
32	34	52.3	431	10	US-09-925-302-485	Sequence 10457, A
33	34	52.3	484	10	US-09-815-242-10457	Sequence 2, Appli
34	33.5	51.5	333	10	US-09-825-882-2	Sequence 1, Appli
35	33	50.8	33	9	US-09-782-704A-1	Sequence 58, Appli
36	33	50.8	100	10	US-09-867-550-58	Sequence 2, Appli
37	33	50.8	112	9	US-10-138-516-2	Sequence 4, Appli
38	33	50.8	112	9	US-10-146-130-4	Sequence 3, Appli
39	33	50.8	112	9	US-10-092-934-3	Sequence 566, App
40	33	50.8	192	10	US-09-764-853-566	Sequence 206, App
41	33	50.8	377	9	US-09-978-295A-206	Sequence 206, App
42	33	50.8	377	9	US-09-978-697-206	Sequence 206, App
43	33	50.8	377	9	US-09-978-192A-206	Sequence 206, App
44	33	50.8	377	9	US-09-999-832A-206	Sequence 206, App
45	33	50.8	377	9	US-09-978-189-206	Sequence 206, App

ALIGNMENTS

RESULT 1
US-09-864-761-37459
Sequence 37459, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 37459
;; LENGTH: 35
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC012002.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
;; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 5.3
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.8
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.5
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.9
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.3
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
;; OTHER INFORMATION: SWISSPROT HIT: P57072, EVALU6 6.40e+00
;; OTHER INFORMATION: EST_HUMAN HIT: BE019472.1, EVALU6 1.70e+00
US-09-864-761-37459
```

```
Query Match          58.5%; Score 38; DB 10; Length 35;
Best Local Similarity 77.8%; Pred. No. 5.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 GVLACAPS 12
    ||| |||
DB 2 GVFACCAPS 10
```

```
RESULT 2
US-10-092-154-737
;; Sequence 737, Application US/10092154
;; Publication No. US20030054375A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC009G1
;; CURRENT APPLICATION NUMBER: US/10/092,154
;; CURRENT FILING DATE: 2002-03-07
;; NUMBER OF SEQ ID NOS: 2003
;; Prior Application removed - See File Wrapper or Palm
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 737
;; LENGTH: 63
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (32)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: misc_feature
;; LOCATION: (44)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-737
```

```
Query Match          58.5%; Score 38; DB 9; Length 63;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 GVLACAP 11
    ||| |||
DB 45 GVLACISP 53
```

```
RESULT 3
US-09-764-847-737
;; Sequence 737, Application US/09764847
;; Patent No. US20020132767A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC009
;; CURRENT APPLICATION NUMBER: US/09/764,847
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 2003
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 737
;; LENGTH: 63
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (32)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (44)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-737
```

```
Query Match          58.5%; Score 38; DB 10; Length 63;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 GVLACAP 11
    ||| |||
DB 45 GVLACISP 53
```

```
RESULT 4
US-08-424-550B-83
;; Sequence 83, Application US/08424550B
;; Patent No. US20020119447A1
;; GENERAL INFORMATION:
;; APPLICANT: JOHN N. SIMONS
;; APPLICANT: TAMM J. PILOT-MATIAS
;; APPLICANT: GEORGE J. DAWSON
;; APPLICANT: GEORGE G. SCHLAUDER
;; APPLICANT: SUPRESE M. DESAI
;; APPLICANT: THOMAS P. LEARY
;; APPLICANT: ANTHONY SCOTT MUEHROFF
;; APPLICANT: JAMES C. ERKER
;; APPLICANT: SHERI L. BUIER
;; APPLICANT: ISA K. MUSHAWAR
;; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;; NUMBER OF SEQUENCES: 716
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;; STREET: 100 ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550B
;; FILING DATE:
;; CLASSIFICATION: 435435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
```

REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-83

Query Match 55.4%; Score 36; DB 8; Length 1422;
Best Local Similarity 87.5%; Pred. No. 6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VGVLAACA 10
||||:||||
Db 1783 VGVLSACA 790

RESULT 5
US-09-742-659-5
Sequence 5, Application US/09742659
Patent No. US20010034019A1
GENERAL INFORMATION:
APPLICANT: Hong, Zhi
APPLICANT: Butkiewicz, Nancy J.
APPLICANT: Zhong, Weidong
APPLICANT: Ingravallo, Paul
APPLICANT: Wright-Minogue, Jacquelyn
APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
FILE REFERENCE: ID01116
CURRENT APPLICATION NUMBER: US/09/742,659
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 5
LENGTH: 2862
TYPE: PRT
ORGANISM: GBV-B/HCV
FEATURE:
NAME/KEY: SITE
LOCATION: (945)..(1129)
OTHER INFORMATION: chimeric region
NAME/KEY: SITE
LOCATION: (1579)..(1593)
OTHER INFORMATION: chimeric region
US-09-742-659-5

Query Match 55.4%; Score 36; DB 10; Length 2862;
Best Local Similarity 87.5%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VGVLAACA 10
||||:||||
Db 1785 VGVLSACA 1792

RESULT 6
US-08-424-550B-394
Sequence 394, Application US/08424550B
Patent No. US20020119447A1
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMT J. PILOT-MATTIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER

APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUEHROFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUTK
APPLICANT: ISA K. MUSHAMMAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/424,550B
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 394:
SEQUENCE CHARACTERISTICS:
LENGTH: 2864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-394

Query Match 55.4%; Score 36; DB 8; Length 2864;
Best Local Similarity 87.5%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VGVLAACA 10
||||:||||
Db 1787 VGVLSACA 1794

RESULT 7
US-09-742-659-2
Sequence 2, Application US/09742659
Patent No. US20010034019A1
GENERAL INFORMATION:
APPLICANT: Hong, Zhi
APPLICANT: Butkiewicz, Nancy J.
APPLICANT: Zhong, Weidong
APPLICANT: Ingravallo, Paul
APPLICANT: Wright-Minogue, Jacquelyn
APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
FILE REFERENCE: ID01116
CURRENT APPLICATION NUMBER: US/09/742,659
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2864
TYPE: PRT
ORGANISM: GB virus-B

US-09-742-659-2

Query Match 55.4%; Score 36; DB 10; Length 2864;
Best Local Similarity 87.5%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VGVLAACA 10
|||||
DB 1787 VGVLSACA 1794

RESULT 8

US-09-742-659-6
Sequence 6, Application US/09742659
Patent No. US20010034019A1
GENERAL INFORMATION:

APPLICANT: Hong, Zhi
APPLICANT: Butkiewicz, Nancy J.
APPLICANT: Zhong, Weidong
APPLICANT: Ingravallo, Paul
APPLICANT: Wright-Minogue, Jacquelyn
APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
FILE REFERENCE: ID01116
CURRENT APPLICATION NUMBER: US/09/742,659
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2865
TYPE: PRT
ORGANISM: GBV-B/HCV
FEATURE:
NAME/KEY: SITE
LOCATION: (2275)..(2865)
OTHER INFORMATION: chimeric region
US-09-742-659-6

Query Match 55.4%; Score 36; DB 10; Length 2865;
Best Local Similarity 87.5%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VGVLAACA 10
|||||
DB 1787 VGVLSACA 1794

RESULT 9

US-09-922-261-95
Sequence 95, Application US/09922261
Patent No. US2002011471A1
GENERAL INFORMATION:

APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kasturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
TITLE OF INVENTION: CELL DEATH
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/922,261
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US/09/461,697
PRIOR FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 95

LENGTH: 52
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-261-95

Query Match 53.8%; Score 35; DB 10; Length 52;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 VLAACAPSK 13
|||||
DB 4 VLAACSPSR 12

RESULT 10

US-09-870-162A-1
Sequence 1, Application US/09870162A
Patent No. US20020042118A1
GENERAL INFORMATION:

APPLICANT: Fuchs, Georg
APPLICANT: Breinig, Sabine
TITLE OF INVENTION: Phenol Induced Proteins of Thauera aromatica
FILE REFERENCE: BC1006 US DIV
CURRENT APPLICATION NUMBER: US/09/870,162A
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/516914
PRIOR FILING DATE: 2000-03-01
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 485
TYPE: PRT
ORGANISM: Thauera aromatica
US-09-870-162A-1

Query Match 53.8%; Score 35; DB 10; Length 485;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 VGVLAACAPS 12
|||||
DB 211 LAAMAACAPS 220

RESULT 11
US-10-211-962-99
Sequence 99, Application US/10211962
Publication No. US20030082640A1
GENERAL INFORMATION:

APPLICANT: Hertz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/10/211,962
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: US/09/562,737
PRIOR FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 99
LENGTH: 749
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-99

Query Match 53.8%; Score 35; DB 9; Length 749;
Best Local Similarity 55.6%; Pred. No. 4.6e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 5 VLAACAPSK 13

Db 268 ILASCSPSR 276

|||||:

RESULT 12

US-09-970-711-12

; Sequence 12, Application US/09970711
; Patent No. US20020081279A1

GENERAL INFORMATION:

; APPLICANT: Baker, Adam

; APPLICANT: Cotten, Matthew

; APPLICANT: Chiocca, Susanna

; APPLICANT: Kuizbauer, Robert

; APPLICANT: Schaffner, Gotthold

; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus

; FILE REFERENCE: 0652.1800001

; CURRENT APPLICATION NUMBER: US/09/970,711

; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 09/171,461

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: PCT/EP97/01944

; PRIOR FILING DATE: 1997-04-18

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 942

; TYPE: PRF

; ORGANISM: CELO Virus

; FEATURE: OTHER INFORMATION: Position: 18289..21117 /gene: L3 /product: L3 hexon

US-09-970-711-12

Query Match 53.8%; Score 35; DB 10; Length 942;

Best Local Similarity 50.0%; Pred. No. 5.9e+02; Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VNVGLACAPS 12

||| | | | |

Db 278 VNSGTRACRPN 289

RESULT 13

US-09-944-849-2

; Sequence 2, Application US/09944849
; Patent No. US20020151487A1

GENERAL INFORMATION:

; APPLICANT: Nickoloff, Brian

; APPLICANT: Mele, Lucio

; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMENT

; FILE REFERENCE: 212583

; CURRENT APPLICATION NUMBER: US/09/944,849

; CURRENT FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/229,614

; PRIOR FILING DATE: 2000-08-31

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2444

; TYPE: PRF

; ORGANISM: Artificial Sequence

; FEATURE: NAME/KEY: misc_feature

; LOCATION: (891)..(891)

; OTHER INFORMATION: The 'Xaa' at location 891 stands for Gly, or Ala.

; NAME/KEY: misc_feature

; LOCATION: (1763)..(1763)

; OTHER INFORMATION: The 'Xaa' at location 1763 stands for Gln, Arg, Pro, or Leu.

; NAME/KEY: misc_feature

; LOCATION: (1787)..(1787)

; OTHER INFORMATION: The 'Xaa' at location 1787 stands for Thr, Ala, Pro, or Ser.

; OTHER INFORMATION: Constitutively Active No. US20020151487A1ch-1

US-09-944-849-2

Query Match 53.8%; Score 35; DB 10; Length 2444;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 VLAACAPS 12

||| ||| |

Db 829 VLAACAPS 836

RESULT 14

US-10-097-065-597

; Sequence 597, Application US/10097065
; Publication No. US20030055236A1

GENERAL INFORMATION:

; APPLICANT: Moore, Paul A. et al.

; TITLE OF INVENTION: 110 Human Secreted Proteins

; FILE REFERENCE: P2021P1

; CURRENT APPLICATION NUMBER: US/10/097,065

; CURRENT FILING DATE: 2002-03-14

; PRIOR APPLICATION NUMBER: PCT/US98/27059

; PRIOR FILING DATE: 1998-12-17

; PRIOR APPLICATION NUMBER: 60/070,923

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,007

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,057

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,006

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,369

; PRIOR FILING DATE: 1997-12-19

; PRIOR APPLICATION NUMBER: 60/068,367

; PRIOR FILING DATE: 1997-12-19

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; PRIOR APPLICATION NUMBER: 60/068,169

; PRIOR FILING DATE: 1997-12-19

; PRIOR APPLICATION NUMBER: 60/068,053

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,064

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,054

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,008

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/068,365

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 672

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 597

; LENGTH: 23

; TYPE: PRF

; ORGANISM: Homo sapiens

US-10-097-065-597

Query Match 52.3%; Score 34; DB 9; Length 23;

Best Local Similarity 60.0%; Pred. No. 19; Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 VGVLAACAPS 12

| | | | |

Db 14 IGHAVCTPS 23

RESULT 15

US-10-097-065-596

; Sequence 596, Application US/10097065
; Publication No. US20030055236A1

GENERAL INFORMATION:

; APPLICANT: Moore, Paul A. et al.

; TITLE OF INVENTION: 110 Human Secreted Proteins

; FILE REFERENCE: P2021P1

;/ CURRENT APPLICATION NUMBER: US/10/097,065
;/ CURRENT FILING DATE: 2002-03-14
;/ PRIOR APPLICATION NUMBER: PCT/US98/27059
;/ PRIOR FILING DATE: 1998-12-17
;/ PRIOR APPLICATION NUMBER: 60/070,923
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,007
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,057
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,006
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,369
;/ PRIOR FILING DATE: 1997-12-19
;/ PRIOR APPLICATION NUMBER: 60/068,367
;/ PRIOR FILING DATE: 1997-12-19
;/ PRIOR APPLICATION NUMBER: 60/068,368
;/ PRIOR FILING DATE: 1997-12-19
;/ PRIOR APPLICATION NUMBER: 60/068,169
;/ PRIOR FILING DATE: 1997-12-19
;/ PRIOR APPLICATION NUMBER: 60/068,053
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,064
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,054
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,008
;/ PRIOR FILING DATE: 1997-12-18
;/ PRIOR APPLICATION NUMBER: 60/068,365
;/ PRIOR FILING DATE: 1997-12-19
;/ NUMBER OF SEQ ID NOS: 672
;/ SOFTWARE: Patentl Ver. 2.0
;/ SEQ ID NO: 596
;/ LENGTH: 135
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
US-10-097-065-396

Query Match 52.3%; Score 34; DB 9; Length 135;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 VGVLIACAPS 12
; 1 1 1 1 1
DB 32 IGHIAVCTPS 41

Search completed: May 20, 2003, 18:00:03
Job time : 6.42857 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:04:12 : Search time 44.6044 Seconds
(without alignments)
122.483 Million cell updates/sec

Title: US-09-869-155-19

Perfect score: 225

Sequence: 1 VAMRFLCLPTGPGVAIFG.....GVPWPQFTQSMPTLVVK 41

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002:*

1: /SID52/gcgdata/geneseq/genesqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/genesqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/genesqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/genesqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/genesqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/genesqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/genesqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/genesqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/genesqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/genesqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/genesqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/genesqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/genesqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/genesqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/genesqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/genesqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/genesqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/genesqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/genesqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/genesqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/genesqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/genesqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/genesqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	41	21	AAV93762
2	225	100.0	41	22	AAEL4683
3	225	100.0	41	22	AAU07398
4	211	93.8	370	23	AAU75816
5	211	93.8	381	23	AAU75813
6	208	92.4	381	23	AAU75812
7	206	91.6	381	23	AAU75814
8	201	89.3	370	23	AAU75815
9	195	86.7	277	23	AAU75818
10	175	77.8	185	23	AAU75824

11	175	77.8	185	23	AAU75825	Barley L endoxylan
12	175	77.8	185	23	AAU75826	Barley L endoxylan
13	114	50.7	287	23	AAU75819	Wheat L endoxylan
14	107	47.6	170	23	AAU75827	Oat L endoxylanase
15	63	28.0	693	23	AB571966	Drosophila melanog
16	62.5	27.8	62	22	AAU75773	Human colon cancer
17	62	27.6	291	22	AAU91308	C glutamylase prote
18	60	26.7	318	21	AB40413	Human ORFX ORF177
19	60	26.7	318	23	ABP05683	Human ORFX protein
20	60	26.7	319	22	AAU45823	Propionibacterium
21	59.5	26.4	258	21	AAU42492	Arabidopsis thalia
22	59.5	26.4	258	21	AAU44139	Arabidopsis thalia
23	59.5	26.4	433	21	AAU42491	Arabidopsis thalia
24	59.5	26.4	433	21	AAU44138	Arabidopsis thalia
25	59.5	26.4	438	21	AAU42490	Arabidopsis thalia
26	59.5	26.4	439	21	AAU44137	Arabidopsis thalia
27	58	25.8	92	22	AAU13472	Human polypeptide
28	57.5	25.6	131	22	ABG12731	Novel human diapo
29	57.5	25.6	169	22	ABG12730	Novel human diapo
30	57	25.3	263	22	ABG92802	Human protein sequ
31	57	25.3	279	22	ABE01202	Human gene 8 encod
32	57	25.3	279	23	ABG64043	Human albumin fusi
33	56.5	25.1	2114	22	ABG21256	Novel human diapo
34	56	24.9	74	22	AAU84155	Human immune/haema
35	56	24.9	119	22	AAU30896	Novel human secret
36	56	24.9	196	23	AAU75817	Wheat L endoxylan
37	55.5	24.7	933	22	ABG19701	Novel human diapo
38	54.5	24.2	142	22	ABG12065	Novel human diapo
39	54.5	24.2	405	21	AAU44494	Arabidopsis thalia
40	54.5	24.2	425	21	AAU44493	Arabidopsis thalia
41	54	24.0	73	22	ABG16583	Human nervous syst
42	54	24.0	95	22	ABG39802	Propionibacterium
43	54	24.0	1202	22	ABG99493	Preliminary human
44	54	24.0	1221	23	ABG61684	Cadherin-like asy
45	54	24.0	1221	23	ABG61692	Cadherin-like asy

ALIGNMENTS

RESULT 1
AAV93762
ID AAV93762 standard; peptide: 41 AA.
AC AAV93762;
DT 03-OCT-2000 (first entry)
DE Amino acid sequence of an endo-beta-1,4-xylanase inhibitor peptide.
XX
XX Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
XX dough; dough preparation.
XX
XX Triflicum sp.
XX
XX WO200039289-A2.
XX
XX 06-JUL-2000.
XX
XX 17-DEC-1999; 99WO-IB02071;
XX
XX 23-DEC-1998; 98GB-0028599.
XX 06-APR-1999; 99GB-0007805.
XX 15-APR-1999; 99GB-0008645.
XX
XX (DANI-) DANISCO AS.
XX
XX Sibbesen O, Sorensen JF;
XX
XX WPI; 2000-465744/40.
XX
XX Mutant xylanase protein identified using xylanase inhibitor useful for
XX preparing non-sticky dough for bakery products -
XX

XX Claim 24; Page 112; 112pp; English.
PS
XX
CC The present sequence is derived from an endo-beta-1,4-xylanase
CC inhibitor. The protein is obtained from wheat flour. The specification
CC also describes a mutant xylanase protein. The xylanase is useful for
CC preparing a foodstuff, preferably a bakery product or a substance
CC (e.g. a dough) for making the bakery product. Wild type xylanase or
CC mutant xylanase is useful for preparing a dough that is less sticky
CC than a dough comprising a fungal xylanase. The xylanase inhibitor is
CC useful for screening high degree resistance xylanases for dough
CC preparation. The xylanase is also useful for preparing a non-sticky
CC dough. A combination of xylanase and the inhibitor is useful for
CC calibrating and/or determining the quantity of inhibitor in a wheat
CC flour sample.
XX
SQ Sequence 41 AA:
Query Match 100.0%; Score 225; DB 21; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VANRFLCLPTGGPGVAIFGGGVPWPQFTQSMPTLVVVK 41
DB 1 VANRFLCLPTGGPGVAIFGGGVPWPQFTQSMPTLVVVK 41
RESULT 2
AAEI4693
ID AAEI4693 standard; peptide; 41 AA.
XX
AC AAEI4693;
XX
DT 21-AUG-2002 (first entry)
XX
DE wheat flour xylanase inhibitor B chain Lys-C digested fragment #4.
XX
KW Refrigerated dough; syrruping; arabinoxylan; bakery product; bread;
KW pizza base; cake; biscuit; wheat; flour; xylanase inhibitor.
XX
OS Trifolium aestivum.
XX
PN WO200152657-A1.
XX
PD 26-JUL-2001.
XX
PE 17-JAN-2001; 2001WO-1B00168.
XX
PR 18-JAN-2000; 2000GB-0001136.
XX
PA (DANI-) DANISCO AS.
XX
PI Poulsen CH, Sorensen JF;
XX
DR WPI; 2001-457446/49.
XX
PT Production of refrigerated dough with reduced syrruping, useful in
PT production of bakery products such as bread, comprises admixing cereal
PT flour, water and protein that prevents enzymatic degradation of
PT arabinoxylan in the cereal flour -
XX
PS Disclosure; Page 23; 26pp; English.
XX
CC The invention relates to a process for producing refrigerated dough
CC with reduced 'syrruping' (precipitation of liquid on the dough surface
CC because of a reduction in water holding capacity caused by the breakdown
CC of arabinoxylan over time). The process comprises admixing cereal flour
CC and water with a protein that reduces/prevents enzymatic degradation of
CC arabinoxylan in the cereal flour. The preferred protein is a xylanase
CC inhibitor. The method is useful to produce refrigerated dough in which
CC syrruping is reduced or eliminated. Refrigerated dough is typically
CC stored for long periods to enable fresh baked products (e.g. bread,
CC pizza bases, cakes, biscuits) to be produced as required. It may also be

CC possible to produce dough or baked products tailored to specific
CC requirements using the method by the use of specific proteins/protein
CC combinations. The present sequence is wheat flour
CC endo-beta-1,4-xylanase inhibitor B chain Lys-C digested fragment.
XX
SQ Sequence 41 AA:
Query Match 100.0%; Score 225; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VANRFLCLPTGGPGVAIFGGGVPWPQFTQSMPTLVVVK 41
DB 1 VANRFLCLPTGGPGVAIFGGGVPWPQFTQSMPTLVVVK 41
RESULT 3
AAU07398
ID AAU07398 standard; protein; 41 AA.
XX
AC AAU07398;
XX
DT 18-DEC-2001 (first entry)
XX
DE Bacillus subtilis xylanase inhibitor #7.
XX
KW Xylanase: plant cell wall; baking; cereal; starch production; wood;
KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; inhibitor.
XX
OS Bacillus subtilis.
XX
PN WO200166711-A1.
XX
PD 13-SEP-2001.
XX
PE 08-MAR-2001; 2001WO-1B00426.
XX
PR 08-MAR-2000; 2000GB-0005585.
XX
PR 27-JUN-2000; 2000GB-0015751.
XX
PA (DANI-) DANISCO AS.
XX
PI Sibsen O, Sorensen JF;
XX
DR WPI; 2001-596834/67.
XX
PT Novel variant xylanase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xylanase inhibitor -
XX
PS Disclosure; Page 63; 70pp; English.
XX
CC The invention relates to a variant xylanase polypeptide (I) or its
CC fragment having xylanase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xylanase inhibitor as compared with the parent xylanase enzyme. (I) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I)
CC is useful for altering the viscosity derived from the presence of
CC hemicellulose or arabinoxylan in a solution or system comprising plant
CC cell wall material. (I) is useful for preparing a foodstuff such as
CC bread, pretzels, tortillas, cakes, cookies, biscuits or crackers.
CC The present sequence represents the amino acid sequence of Bacillus
CC subtilis xylanase inhibitor #7 as described in the method of the
CC invention.
XX
SQ Sequence 41 AA:
Query Match 100.0%; Score 225; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VARNFLCLPTGGGVAIFGGGVPWPQFQSMPTLVVVK 41
 |||
 Db 1 VARNFLCLPTGGGVAIFGGGVPWPQFQSMPTLVVVK 41

RESULT 4
 AAU75816
 ID AAU75816 standard; Protein; 370 AA.
 XX
 AC AAU75816;
 XX
 DT 23-APR-2002 (first entry)
 XX

Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.02.
 DE
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 XX Triticum aestivum cultivar Estica.
 OS
 XX MO200198474-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX

(LEUV-) LEUVEN RES & DEV.
 PA
 XX Delcour J, Debyser W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Campenhout S;
 PI
 DR MPI; 2002-114579/15.
 DR N-PSDB; ABK13673.
 DR

Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 PT
 XX
 PS Claim 127; Page 57; 127pp: English.

The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or macroorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable

CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-I ('T. aestivum L endoxylanase
 CC inhibitor').
 CC
 XX
 SQ Sequence 370 AA;

Query Match 93.8%; Score 211; DB 23; Length 370;
 Best Local Similarity 92.7%; Pred. No. 2, 2e-19;
 Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VARNFLCLPTGGGVAIFGGGVPWPQFQSMPTLVVVK 41
 |||
 Db 149 VARNFLCLPTGGGVAIFGGGVPWPQFQSMPTLVVVK 189

RESULT 5
 AAU75813
 ID AAU75813 standard; Protein; 381 AA.
 XX
 AC AAU75813;
 XX
 DT 23-APR-2002 (first entry)
 XX

Wheat L endoxylanase inhibitor, TAXI I, variant #2.
 DE
 XX Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 KM
 XX Triticum aestivum.
 OS
 XX
 FH Key Location/Qualifiers
 FH Misc-difference 95
 FT /label= Unknown
 FT Misc-difference 98
 FT /label= Unknown
 FT Misc-difference 101
 FT /label= Unknown
 FT Misc-difference 110
 FT /label= Unknown
 FT Misc-difference 333
 FT /label= Unknown
 PN WO200198474-A1.
 XX
 XX 27-DEC-2001.
 PD
 XX 21-JUN-2001; 2001WO-BE00106.
 PF
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX

(LEUV-) LEUVEN RES & DEV.
 PA
 XX Delcour J, Debyser W, Gebruers K, Goesaert H, Flerens K, Robben J;
 PI Van Campenhout S;
 PI
 DR MPI; 2002-114579/15.

SQ Sequence 381 AA:
 Query Match 92.4%; Score 208; DB 23; Length 381;
 Best Local Similarity 90.2%; Pred. No. 5.6e-19;
 Matches 37; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VANRFLCLPTGGPVAIFGGGVPWPQFTQSMPTLVYK 41
 ||||||||||||||||||||||||||||||||
 DB 160 VANRFLCLPTGGPVAIFGGGVPWPQFTQSMPTLVYK 200

RESULT 7
 AAU75814
 ID AAU75814 standard; Protein; 381 AA.
 XX AAU75814;
 AC 23-APR-2002 (first entry)
 DT
 XX
 DE Wheat L endoxylanase inhibitor, TAXI I, microheterogenic variant.
 XX
 KW Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Triticum aestivum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95 /label= Unknown
 FT Misc-difference 98 /label= Unknown
 FT Misc-difference 101 /label= Unknown
 FT Misc-difference 110 /label= Unknown
 FT Misc-difference 145 /label= Unknown
 FT Misc-difference 183 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 232 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 275 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 282 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 333 /label= Asp, Asn, Val, Leu, Ser, Phe, Pro, Ala, Cys
 FT Misc-difference 333 /label= Unknown
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-BE00106.
 XX
 PR 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEEVEN RES & DEV.
 XX
 PI Delcours J, Debysier W, Gebruers K, Goesaert H, Fierens K, Robben J;
 XX Van Campenhout S;
 DR WPI; 2002-114579/15.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during

PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 51; 127pp; English.
 XX
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syripping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a microheterogenic variant of wheat TAXI-I (T. aestivum L
 CC endoxylanase inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 SQ Sequence 381 AA:
 Query Match 91.6%; Score 206; DB 23; Length 381;
 Best Local Similarity 90.2%; Pred. No. 1e-18;
 Matches 37; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VANRFLCLPTGGPVAIFGGGVPWPQFTQSMPTLVYK 41
 ||||||||||||||||||||||||||||||||
 DB 160 VANRFLCLPTGGPVAIFGGGVPWPQFTQSMPTLVYK 200

RESULT 8
 AAU75815
 ID AAU75815 standard; Protein; 370 AA.
 XX
 AC AAU75815;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Wheat L endoxylanase inhibitor, TAXI I, partial sequence TAXI-I.0.
 XX
 KW Wheat; TAXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Triticum aestivum cultivar Soissons.
 XX
 PN WO200198474-A1.
 XX
 PD 27-DEC-2001.

[illegible]

XX	DT	23-APR-2002	(first entry)
XX	DE	L endoxylanase inhibitor, ATRXI-II, partial sequence ATRXI-II.01.	
XX	XX		
XX	XX	ATRXI-II: L endoxylanase inhibitor; cellulolytic enzyme inhibitor; ds: xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor; ds: immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant; straw; light dough; sponge dough; Chorleywood bread; biscuits; pasta; noodle; animal feed; starch separation; maize processing; malting; plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.	
XX	XX		
XX	OS	Aegilops tauschii variety Acc2220051.	
XX	XX		
XX	XX	Key	
XX	XX	Location/Qualifiers	
XX	XX	Misc-difference 4	
XX	XX	/note= "Encoded by CTC"	
XX	XX		
XX	XX	27-DEC-2001.	
XX	XX		
XX	XX	21-JUN-2001; 2001WO-BE00106.	
XX	XX		
XX	XX	22-JUN-2000; 2000GB-0015296.	
XX	XX	25-JAN-2001; 2001GB-0002018.	
XX	XX	26-JAN-2001; 2001GB-0002194.	
XX	XX	16-MAR-2001; 2001GB-0006564.	
XX	XX	21-MAY-2001; 2001GB-0012328.	
XX	XX		
XX	PA	(LEUV-) LEUVEN RES & DEV.	
XX	PI		
XX	PI	Delcourt J, Debyser W, Gebruers K, Goesaert H, Fierens K, Robben J; Van Campenhout S;	
XX	XX		
XX	DR	WPI: 2002-114579/15.	
XX	DR	N-PSDB; ABRK13675.	
XX	XX		
XX	PT	Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or beta-glucanolytic enzymes comprises using endoxylanases during screening for inhibition activity or affinity chromatography with immobilised enzymes	
XX	PT		
XX	PS	Claim 127; Page 58; 127pp; English.	
XX	XX		
XX	XX	The invention relates to separating and/or isolating inhibitors of cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulose, xylan, arabinoxylan or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and modulators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylanases that are totally, less or not inhibited by the inhibitors, reducing syruping in rehydrated dough compositions, affecting the relative affinity and/or relative hydrolysis specificity and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylanases such as by the formation of an endoxylanase/inhibitor complex, improving the malting of cereals such as barley, sorghum and wheat and the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten starch separation and production, improving maize processing,	

CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial Aegilops tauschii HVXI-II (Aegilops
 CC tauschii L endoxylanase inhibitor).

XX Sequence 277 AA:

Query Match 86.7%: Score 195; DB 23; Length 277;
 Best Local Similarity 82.9%: Pred. No. 2e-17;
 Matches 34; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 VANRFLCLPTGGVAFIFGGPVPMPQFTQSMPTLVVVK 41
 DB 149 VANKFLCLPISGGVAFIFGGPPLPMPQLTQSMPTLVVVK 189

RESULT 10

AAU75824
 ID AAU75824 standard; Protein; 185 AA.

XX AAU75824;

XX 23-APR-2002 (first entry)

DE Barley L endoxylanase inhibitor, HVXI I, variant #1.

XX Barley: HVXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;

KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;

KM immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;

KM straight dough; sponge dough; Chorleywood bread; biscuits; pasta;

KM noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX Hordeum vulgare.

XX Hordeum vulgare.

XX WO200198474-A1.

XX 27-DEC-2001.

XX 21-JUN-2001; 2001WO-BE00106.

XX 22-JUN-2000; 2000GB-0015296.

XX 25-JAN-2001; 2001GB-0002018.

XX 26-JAN-2001; 2001GB-0002194.

XX 16-MAR-2001; 2001GB-0006564.

XX 21-MAY-2001; 2001GB-0012328.

XX (LEUV-) LEUVEN RES 6 DEV.

XX Delcourt J, Debysier W, Gebruers K, Goesaert H, Flerens K, Robben J;

XX Van Campenhout S;

XX WPI; 2002-114579/15.

XX Claim 127; Page 63; 127pp; English.

CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also

CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syripping in rehydrated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of barley HVXI-I (H. vulgare L endoxylanase
 CC inhibitor).

CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

XX Sequence 185 AA:

Query Match 77.8%: Score 175; DB 23; Length 185;
 Best Local Similarity 78.0%: Pred. No. 5.4e-15;
 Matches 32; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 VANRFLCLPTGGVAFIFGGPVPMPQFTQSMPTLVVVK 41
 DB 20 VSHRFLCLPTGGVAFIFGGPPLPMPQLTQSMPTLVVVK 60

RESULT 11

AAU75825
 ID AAU75825 standard; Protein; 185 AA.

XX AAU75825;

XX 23-APR-2002 (first entry)

DE Barley L endoxylanase inhibitor, HVXI I, variant #2.

XX Barley: HVXI-I: L endoxylanase inhibitor; cellulolytic enzyme inhibitor;

KM xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;

KM immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;

KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;

KW noodle; animal feed; starch separation; maize processing; malting;

KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

XX Hordeum vulgare.

XX Hordeum vulgare.

XX WO200198474-A1.

XX 27-DEC-2001.

XX 21-JUN-2001; 2001WO-BE00106.

XX 22-JUN-2000; 2000GB-0015296.

XX 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcour J, Debyser W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI; 2002-114579/15.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 PS Claim 127; Page 63; 127pp: English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and pharmaceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a variant of barley HvXI-1 (H. vulgare L endoxylanase
 CC inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.
 XX
 SQ Sequence 185 AA:
 Query Match 77.8%; Score 175; DB 23; Length 185;
 Best Local Similarity 78.0%; Pred. No. 5.4e-15;
 Matches 32; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VARNRLCLPTGGPGVAIFGGGVPMPOFTQSMPTLVVVK 41
 DB 20 VSHRFLCLPTGGAGVAILGGGLPMPQFTQSMAYTPLVGK 60
 RESULT 12
 AAU75326
 ID AAU75826 standard; Protein; 185 AA.
 XX
 AC AAU75826;
 XX
 XX 23-APR-2002 (first entry)
 DT
 XX
 DE Barley L endoxylanase inhibitor, HvXI 1, microheterogenic variant.

XX
 KW Barley; HvXI-1; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syruping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.
 XX
 OS Hordeum vulgare.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FH MISC-difference 72
 FT /label= Unknown
 FT MISC-difference 128
 FT /label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
 FT MISC-difference 132
 FT /label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
 FT MISC-difference 134
 FT /label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
 FT MISC-difference 148
 FT /label= Pro, Ala, Asp, Leu, Gly, Asn, Cys
 FT MISC-difference 165
 FT /label= Unknown
 PN
 XX
 PD WO200198474-A1.
 XX
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-BE0106.
 XX
 XX 22-JUN-2000; 2000GB-0015296.
 PR 25-JAN-2001; 2001GB-0002018.
 PR 26-JAN-2001; 2001GB-0002194.
 PR 16-MAR-2001; 2001GB-0006564.
 PR 21-MAY-2001; 2001GB-0012328.
 XX
 PA (LEUV-) LEUVEN RES & DEV.
 PI Delcour J, Debyser W, Gebruers K, Goesaert H, Fierens K, Robben J;
 PI Van Campenhout S;
 DR WPI; 2002-114579/15.
 XX
 PT Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or
 PT beta-glucanolytic enzymes comprises using endoxylanases during
 PT screening for inhibition activity or affinity chromatography with
 PT immobilised enzymes
 XX
 XX Claim 127; Page 63; 127pp: English.
 CC The invention relates to separating and/or isolating inhibitors of
 CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprises
 CC screening the inhibition activity by using two or more enzymes during the
 CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoxylan or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syruping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinoxylans of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as

CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence is a microheterogenic variant of barley HXVI-1 (H. vulgare L
 CC endoxylanase inhibitor).
 CC Note: Variant amino acids are highlighted in the specification but
 CC no wild-type sequence is shown for comparison.

SQ Sequence 185 AA:

Query Match 77.8% Score 175; DB 23; Length 185;
 Best Local Similarity 78.0%; Pred. No. 5.4e-15;
 Matches 32; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 VANRFLCLPTGGVAIFGSGVPMWPFOTQSMPTLVVVK 41
 Db 20 VSHRFLCLPTGGAGVAILLGGPLPMPFTQSMATPVLVK 60

RESULT 13
 AAU75819
 ID AAU75819 standard; Protein: 287 AA.

AC AAU75819;

DT 23-APR-2002 (first entry)

DE Wheat L endoxylanase inhibitor, partial sequence TAXI-III.

KW Wheat, TAXI-III; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Triticum aestivum cultivar Solissons.

PN WO200198474-A1.

PD 27-DEC-2001.

PF 21-JUN-2001; 2001WO-BE00106.

PR 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.

PR 16-MAR-2001; 2001GB-0006564.

PR 21-MAY-2001; 2001GB-0012328.

PA (LEUV-) LEUVEN RES & DEV.

PI Delcour J, Debysse W, Gebruers K, Goesaert H, Piereens K, Robben J;

PI Van Campenhout S;

DR WPI; 2002-114579/15.

DR N-PSDB; ABR13676.

XX Separating and/or isolating inhibitors of cellulolytic, xylanolytic, or

PT beta-glucanolytic enzymes comprising endoxylanases during

PT screening for inhibition activity or affinity chromatography with

PT immobilised enzymes

XX Claim 127; Page 58; 127pp; English.

PS The invention relates to separating and/or isolating inhibitors of

CC cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising

CC screening the inhibition activity by using two or more enzymes during the

CC separation and/or isolation steps that allow to distinguish inhibitors of
 CC different specificity or by using an affinity chromatographic step with
 CC immobilised enzymes and/or antibodies against inhibitors. Also
 CC included are an isolated nucleic acid molecule encoding an inhibitor
 CC which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase,
 CC alpha-L-arabino-furanosidase and/or other cellulose, xylan,
 CC arabinoside or beta-glucan degrading enzymes, a vector comprising the
 CC nucleic acid, an expression system transformed with the nucleic
 CC acid, a host organism transformed with the nucleic acid, the inhibitory
 CC proteins encoded by the nucleic acids and modulators of the proteins.
 CC A recombinant protein, glycoprotein or polypeptide or microorganisms,
 CC plant or plant materials transformed with the nucleic acid are useful
 CC for the formation of an endoxylanase-inhibitor complex, screening
 CC endoxylanases that are totally, less or not inhibited by the inhibitors,
 CC reducing syripping in refrigerated dough compositions, affecting the
 CC relative affinity and/or relative hydrolysis specificity and/or relative
 CC hydrolysis rate versus water-extractable and/or water-unextractable
 CC arabinosides of endoxylanases such as by the formation of an
 CC endoxylanase/inhibitor complex, improving the malting of cereals such as
 CC barley, sorghum and wheat and/or the production of beer, improving the
 CC production and/or quality of baked or extruded cereal products such as
 CC straight dough, sponge dough, Chorleywood bread, breakfast cereals,
 CC biscuits, pasta and noodles, animal feed stuff, improving the production
 CC of starch derived syrups, sorbitol, xylitol, xylose and/or xylitol, wheat gluten
 CC starch separation and production, improving maize processing,
 CC plant disease resistance and nutraceutical and/or pharmaceutical
 CC applications, improving paper and pulp technologies. The present
 CC sequence represents a partial wheat TAXI-III (T. aestivum L endoxylanase
 CC inhibitor).

SQ Sequence 287 AA:

Query Match 50.7% Score 114; DB 23; Length 287;
 Best Local Similarity 49.0%; Pred. No. 7.9e-07;
 Matches 24; Conservative 8; Mismatches 9; Indels 8; Gaps 2;

OY 1 VANRFLCLPTGG--GPGVAIFGSGPV----PMPFTQSMPTLVVVK 41
 Db 151 VSNKFLCLPTGRGSSDPGVAIFGSGPLHFMARREDTRELATPVLAK 199

RESULT 14

AAU75827
 ID AAU75827 standard; Protein: 170 AA.

AC AAU75827;

DT 23-APR-2002 (first entry)

DE Oat L endoxylanase inhibitor, partial sequence ASXI-01.

KW Oat; ASXI-I; L endoxylanase inhibitor; cellulolytic enzyme inhibitor;
 KW xylanolytic enzyme inhibitor; beta-glucanolytic enzyme inhibitor;
 KW immobilised enzyme; enzyme; dough syripping; cereal product; beer; plant;
 KW straight dough; sponge dough; Chorleywood bread; biscuits; pasta;
 KW noodle; animal feed; starch separation; maize processing; malting;
 KW plant disease resistance; nutraceutical; pharmaceutical; paper; pulp.

OS Avena sativa.

PN WO200198474-A1.

PD 27-DEC-2001.

PF 21-JUN-2001; 2001WO-BE00106.

PR 22-JUN-2000; 2000GB-0015296.

PR 25-JAN-2001; 2001GB-0002018.

PR 26-JAN-2001; 2001GB-0002194.

PR 16-MAR-2001; 2001GB-0006564.

PR 21-MAY-2001; 2001GB-0012328.

PA (LEUV-) LEUVEN RES & DEV.

Pt
Pt separating and/or isolating inhibitors of cellulolytic, xylanolytic, or beta-glucanolytic enzymes comprises using endoxylases during screening for inhibition activity or affinity chromatography with immobilised enzymes

Xx Claim 127; Page 65; 127/p; English.

Cc The invention relates to separating and/or isolating inhibitors of cellulolytic, xylanolytic and/or beta-glucanolytic enzymes comprising screening the inhibition activity by using two or more enzymes during the separation and/or isolation steps that allow to distinguish inhibitors of different specificity or by using an affinity chromatographic step with immobilised enzymes and/or antibodies against inhibitors. Also included are an isolated nucleic acid molecule encoding an inhibitor which inhibits cellulase, endoxylanase, beta-glucanase, beta-xylosidase, alpha-L-arabino-furanosidase and/or other cellulose, xylan, arabinoxyran or beta-glucan degrading enzymes, a vector comprising the nucleic acid, an expression system transformed with the nucleic acid, a host organism transformed with the nucleic acid, the inhibitory proteins encoded by the nucleic acids and mediators of the proteins. A recombinant protein, glycoprotein or polypeptide or microorganisms, plant or plant materials transformed with the nucleic acid are useful for the formation of an endoxylanase-inhibitor complex, screening endoxylases that are totally, less or not inhibited by the inhibitors, reducing syruiping in refrigerated dough compositions, affecting the relative affinity and/or relative hydrolysis specifically and/or relative hydrolysis rate versus water-extractable and/or water-unextractable arabinoxylans of endoxylases such as by the formation of an endoxylanase/inhibitor complex, improving the malting of cereals such as barley, sorghum and wheat and/or the production of beer, improving the production and/or quality of baked or extruded cereal products such as straight dough, sponge dough, Chorleywood bread, breakfast cereals, biscuits, pasta and noodles, animal feed stuff, improving the production of starch derived syrups, sorbitol, xylose and/or xyllitol, wheat gluten starch separation and production, improving maize processing, CC plant disease resistance and nutraceutical and/or pharmaceutical applications, improving paper and pulp technologies. The present sequence represents a partial oat ASX-I (A. sativa L endoxylanase inhibitor).

Sq Sequence 170 AA;

Oy Query Match . 47.6%; Score 107; DB 23; Length 170;
Best local Similarity 51.0%; Pared No. 3.7e-06;
Matches 25; Conservative 5; Mismatched 11; Indels 8; Gaps 2

Dy 1 VANNRLCLPLTNG---GPGVAIFGGGPV-----PWQFQTSMPTLVVK 41
||| |||| | | |||||||| : |||| | | :
Db 6 VAKRKLGLLSRGVGYGDAVFIRGGPILHITNQPEPDYDQSLEYTLFTK 54

Rf RESULT 15
ABBT1966
ID ID T ABBT1966 standard; Protein; 693 AA.
XX AC / ABB71966:
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 42690.
KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.
XS Drosophila melanogaster.

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:11:22 ; Search time 16.6703 seconds
(without alignments)
236.439 Million cell updates/sec

Title: US-09-869-155-19

Perfect score: 225

Sequence: 1 VANNFLCLPTGPGGVAIFG.....GPVPMPOFTOSMPYTLVVK 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62.5	27.8	274	1 D70378	conserved hypotet
2	62	27.6	301	2 T36884	hypothetical prote
3	61	27.1	446	2 G87448	succinylarginine d
4	59.5	26.4	301	2 A99940	hypothetical prote
5	59.5	26.4	301	2 E85788	hypothetical prote
6	59.5	26.4	301	2 D64938	hypothetical prote
7	59.5	26.4	433	2 F86163	hypothetical prote
8	59.5	26.4	534	2 A99316	hypothetical prote
9	59	26.2	238	2 G64868	cation transport p
10	59	26.2	307	2 F70502	hypothetical prote
11	59	26.2	311	2 S72967	hypothetical prote
12	59	26.2	434	2 G86163	hypothetical prote
13	58	25.8	216	2 F70073	hypothetical prote
14	57	25.3	127	2 T38856	hypothetical prote
15	57	25.3	351	2 I46033	NIPP-1, nuclear in
16	56.5	25.1	1060	2 T31763	hypothetical prote
17	56	24.9	291	2 AH0262	conserved hypotet
18	56	24.9	305	2 S70183	chemotaxis protein
19	55.5	24.7	1050	2 T31853	hypothetical prote
20	55	24.4	629	2 E47096	membrane transloc
21	54	24.0	388	2 DB2578	conserved hypotet
22	53.5	23.8	325	2 D82662	hypothetical prote
23	53.5	23.8	684	2 A56154	Abl substrate ena
24	53.5	23.8	768	2 T35465	hypothetical plasm
25	53	23.6	314	2 T03524	codb protein - Rho
26	53	23.6	314	2 S52223	hypothetical prote
27	53	23.6	361	2 T12470	hypothetical prote
28	53	23.6	699	1 I54763	Ra-reactive factor
29	52.5	23.3	356	2 T19792	hypothetical prote

30	52.5	23.3	505	1 S38534	cytochrome P450 76
31	52.5	23.3	1421	2 T05892	hypothetical prote
32	52	23.1	671	2 T36037	probable export as
33	52	23.1	1360	2 T31674	hypothetical prote
34	52	23.1	1410	2 T15661	hypothetical prote
35	51.5	22.9	179	2 B64861	yele protein - Esc
36	51.5	22.9	540	2 S54586	probable membrane
37	51.5	22.9	1334	2 T50568	probable multi-dom
38	51.5	22.9	3190	2 T13828	CREB-binding prote
39	51	22.7	231	2 B65702	cation transport r
40	51	22.7	239	2 C90844	cation transport r
41	51	22.7	295	2 D83259	conserved hypotet
42	51	22.7	324	2 C69261	2-hydroxyhepta-2,4
43	51	22.7	332	2 F75473	conserved hypotet
44	51	22.7	361	2 T38638	hypothetical prote
45	51	22.7	452	2 A57217	type protein - fru

ALIGNMENTS

```

RESULT 1
D70378
conserved hypothetical protein aq_909 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: D70378
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: D70378
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-274 <AOP>
A:Cross-references: GB:AE000713; GB:AE000657; NID:g2983424; PIDN:AAC07026.1; PID:g298
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_909
C:Superfamily: conserved hypothetical protein H10072

Query Match
Best Local Similarity 27.8%; Score 62.5; DB 1; Length 274;
Matches 13; Conservative 6; Mismatches 13; Indels 5; Gaps 1;

QY 6 LCLPTGPGVAIFGGPVPMPO----FTQSMPTL 37
:| | | | | :| | | | | :| |
Db 166 ILSTPGSTAVALSAGGPVPESSQLPLVPCPHL 202

RESULT 2
T36884
hypothetical protein SC151.21c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T36884
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21617
A:Accession: T36884
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-301 <MUR>
A:Cross-references: EMBL:AL109848; PIDN:CAE52845.1; GSPDB:GN00070; SCOEDB:SC151.21c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC151.21c
C:Superfamily: conserved hypothetical protein H10072

Query Match
Best Local Similarity 27.6%; Score 62; DB 2; Length 301;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

```

```

OY      6  LILCLPTGPGVAIFGGPVPWPQ 28
      :| ||| | ||| |||
      184 VLSTPTGTAATATAGCPVWPWE 206

RESULT 3
G87443
succinylarginine dihydrolase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: G87448
R:Niezman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Taub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: G87448
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-446 <STO>
A:Cross-references: GB:AE005673; NID:g13423007; PIDN:AAK23587.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1608
C:Superfamily: Escherichia coli hypothetical protein b1745

Query Match
Best Local Similarity 27.1%; Score 61; DB 2; Length 446;
Matches 13; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

OY      2  ANRFLCLPTGPGVAIFGGPVPWPQFTQSM 34
      :| ||| ||| | | | |
      178 ANHVRCAEHGCGVNLFWGREGAMSHMGCRF 210

Db      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

RESULT 4
A99940
hypothetical protein ECS2489 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: A99940
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A99940
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-301 <HAV>
A:Cross-references: GB:BA000007; PIDN:BAH35912.1; PID:g13361956; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECS2489
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D16.200

Query Match
Best Local Similarity 26.4%; Score 59.5; DB 2; Length 301;
Matches 14; Conservative 4; Mismatches 5; Indels 13; Gaps 1;

OY      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

Db      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

RESULT 5
E85788
hypothetical protein yead [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E85788
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

```

```

Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimlantia, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: E85788
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-301 <STO>
A:Cross-references: GB:AE005174; NID:g12515814; PIDN:AA656769.1; GSPDB:GN00145; UNCP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yead
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D16.200

Query Match
Best Local Similarity 26.4%; Score 59.5; DB 2; Length 301;
Matches 14; Conservative 4; Mismatches 5; Indels 13; Gaps 1;

OY      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

Db      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

RESULT 6
D64938
hypothetical protein b1780 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: D64938
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64938
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-301 <BLAT>
A:Cross-references: GB:AE000273; GB:U00096; NID:g1788078; PIDN:AAK74850.1; PID:g17880
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D16.200

Query Match
Best Local Similarity 26.4%; Score 59.5; DB 2; Length 301;
Matches 14; Conservative 4; Mismatches 5; Indels 13; Gaps 1;

OY      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

Db      15  GVAIFGCGVPWP-----QFTQSMPTL 37
      :| ||| ||| | | | |
      77 GVAIRGCVPCWPCWPGPAOQGLPAHGFAARNLPWL 112

RESULT 7
F66163
hypothetical protein F15K9.17 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: F66163
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Keul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: F66163
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <STO>
A:Cross-references: GB:AE005172; NID:g3850579; PIDN:AAK72119.1; GSPDB:GN00141

```

C:Genetics:
A:Map position: 1
C:Superfamily: conglutin gamma

Query Match 26.4%; Score 59.5; DB 2: Length 433;
Best Local Similarity 60.0%; Pred. No. 6.8;
Matches 12; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 4 RFLCPTGGPGVAIFGGP 23
| : || | | | | | | | | |
DB 199 KFAVCL-TSGKVAFFGNGP 217

RESULT 8

A99316
hypothetical protein norB-1 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: A99316

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Cost, J.
Submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: A99316

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1534 <KUR>

A:Cross-references: GB:AE006641; NID:g13814802; PIDN:AAK41784.1; GSPDB:GN00155

C:Genetics:

A:Gene: norB-1

Query Match 26.4%; Score 59.5; DB 2: Length 534;
Best Local Similarity 38.7%; Pred. No. 8.4;
Matches 12; Conservative 7; Mismatches 9; Indels 3; Gaps 2;

QY 13 GPGV-AIFGGGPV--PWPQFTQSMPTLVVV 40
| : || | | | | : | | | | : : : :
DB 240 GSGGAIFFGSLPIAPMPNFTEDQFLIMIM 270

RESULT 9

G64868
cation transport protein chac - Escherichia coli (strain K-12)

C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: G64868

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Ce-
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MIMD:97426617; PMID:9278503

A:Accession: G64868

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-238 <BLAT>

A:Cross-references: GB:AE000220; GB:U00096; NID:g1787467; PIDN:AAC74302.1; PID:g1787470;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: chac

C:Keywords: transport protein

Query Match 26.2%; Score 59; DB 2: Length 238;
Best Local Similarity 47.2%; Pred. No. 4.3;
Matches 17; Conservative 3; Mismatches 12; Indels 4; Gaps 2;

QY 6 LLCLPTGGPGVAIFGGPVPMP--OFTQSMPTLV 38
| | | | | | | | | | | | | | | | | | | |
DB 45 LACRPDGP-VWIFGYSLMNPALETFESTGTLV 79

RESULT 10

F70502
hypothetical protein RV1695 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70502

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MIMD:98295987; PMID:9634230

A:Accession: F70502

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1307 <COL>

A:Cross-references: GB:Z98268; GB:AL123456; NID:g3261839; PIDN:CAB10952.1; PID:g23267

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV1695

C:Superfamily: conserved hypothetical protein HI0072

Query Match 26.2%; Score 59; DB 2: Length 307;
Best Local Similarity 50.0%; Pred. No. 5.5;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 LLCLPTGGPGVAIFGGPVPMP 27
| : || | | | | | | | | | | | | | | |
DB 192 LVSTPTGSTAVAFSAGGPVLM 213

RESULT 11

S72967
hypothetical protein - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C:Accession: S72967

R:Smith, D.R.; Robison, K.

Submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid L247.

A:Reference number: S72589

A:Accession: S72967

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <SMID>

A:Cross-references: EMBL:U00021; NID:g467141; PIDN:AAA50923.1; PID:g467159

C:Genetics:

A:Start codon: GTG

C:Superfamily: conserved hypothetical protein HI0072

Query Match 26.2%; Score 59; DB 2: Length 311;
Best Local Similarity 50.0%; Pred. No. 5.6;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 LLCLPTGGPGVAIFGGPVPMP 27
| : || | | | | | | | | | | | | | | |
DB 196 LVSTPTGSTAVAFSAGGPVLM 217

RESULT 12

G86163

hypothetical protein F15K9_16 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: G86163

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: AB61441; MUID:21016719; PMID:11130712

A:Accession: G86163

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-434 <STO>

A:Cross-references: GB:AE005172; NID:g3850580; PIDN:AMC72120.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: conglutin gamma

Query Match 26.2%; Score 59; DB 2; Length 434;

Best Local Similarity 38.5%; Pred. No. 7.9;

Matches 15; Conservative 7; Mismatches 13; Indels 4; Gaps 2;

OY 4 RFLCLPTGGGVAIFGCGP---VPMPOFTQSMPTLVV 39

DB 200 KFAVCL-TSGRGVAFGNGPYVFLPGIQISRLQKTPLLI 237

RESULT 13

F70073

hypothetical protein yxce - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: F70073

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Brun, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall

lechi, C.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huijlo, M.F.

Koester, P.; Konungstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lindrois,

A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schreier, R.; Scrofano, F.; Sekiguchi, J.; Sekowska, A.; Serp

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Walters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033; PMID:934377

A:Accession: F70073

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-216 <KUN>

A:Cross-references: GB:299124; GB:AL009126; NID:g2636442; PIDN:CAB16015.1; PID:el184704;

A:Experimental source: strain 168

C:Genetics:

A:Gene: yxce

Query Match 25.8%; Score 58; DB 2; Length 216;

Best Local Similarity 47.1%; Pred. No. 5.2;

Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 10 PTGGPVAIFGCGPVPV 26

DB 169 PVGPGVYSGASPIAW 185

RESULT 14

I38855

activator of RNA decay - human

C:Species: *Homo sapiens* (man)

C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000

C:Accession: I38856

R:Wang, M.; Cohen, S.N.

A:Title: ard-1: a human gene that reverses the effects of temperature-sensitive and dele

A:Reference number: I38856; MUID:95024160; PMID:7524097

A:Accession: I38856

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-127 <RES>

A:Cross-references: EMBL:U014575; NID:g559771; PIDN:AAA64749.1; PID:g559772

C:Genetics:

A:Gene: ard-1

Query Match 25.3%; Score 57; DB 2; Length 127;

Best Local Similarity 40.0%; Pred. No. 4;

Matches 12; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 10 PTGGPVAIFGCGPVPVPMPOFTQSMPTLVV 39

DB 54 PHGHTGALIGLPMYPNLPADVDLTPVV 83

RESULT 15

I46033

NIPP-1, nuclear inhibitor of protein phosphatase-1 - bovine

C:Species: *Bos primigenius* (cattle)

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I46033

R:Van Eynde, A.; Wera, S.; Beullens, M.; Torrekens, S.; van Leuven, F.; Stalmans, W.;

J. Biol. Chem. 270, 28068-28074, 1995

A:Title: Molecular cloning of NIPP-1, a nuclear inhibitor of protein phosphatase-1, r

A:Reference number: I46033; MUID:96081835; PMID:7499293

A:Accession: I46033

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-351 <VAN>

A:Cross-references: EMBL:250748; NID:g1082085; PIDN:CAA90625.1; PID:g1082086

Query Match 25.3%; Score 57; DB 2; Length 351;

Best Local Similarity 40.0%; Pred. No. 11;

Matches 12; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 10 PTGGPVAIFGCGPVPVPMPOFTQSMPTLVV 39

DB 278 PHGHTGALIGLPMYPNLPADVDLTPVV 307

Search completed: May 20, 2003, 16:37:08

Job time: 18.6703 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:05:42 ; Search time 11.0385 Seconds

(without alignments)
154.055 Million cell updates/sec

Title: US-09-869-155-19

Perfect score: 225
Sequence: 1 VANRFLCLPTGCGEVAIFG.....GVPWPQFTQSMPTLVVVK 41

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	28.0	693	1 CAUP_DROME	P54269 drosophila
2	62.5	27.8	274	1 PPNK_AQUAE	O67055 aquifex aeo
3	62	27.6	301	1 PPNK_STRCO	Q9s219 streptomyce
4	59.5	26.4	294	1 UP08_ECOLI	P39173 escherichia
5	59	26.2	231	1 CHAC_ECOLI	P39163 escherichia
6	59	26.2	307	1 PPNK_MYCTU	O33196 mycobacteri
7	59	26.2	311	1 PPNK_MYCTU	O49897 mycobacteri
8	58	25.6	216	1 YXCE_BACSU	P46335 bacillus su
9	57.5	25.3	351	1 PPR8_BOVIN	O28147 bos taurus
10	57	25.3	351	1 PPR8_BOVIN	O12972 homo sapien
11	57	25.3	314	1 COBD_RHOCA	O52683 rhodobacter
12	53	23.6	699	1 CRAR_HUMAN	P48740 h complemen
13	53	23.6	356	1 IRXH_CAEEL	O91644 caenorhabdl
14	52.5	23.3	441	1 PAC1_MOUSE	O61644 mus musculu
15	52.5	23.3	441	1 PAC1_MOUSE	O91644 mus musculu
16	52.5	23.3	441	1 PAC1_MOUSE	O91644 mus musculu
17	52.5	23.3	444	1 PAC1_MOUSE	O91644 mus musculu
18	52.5	23.3	505	1 C762_SOLME	O9b111 homo sapien
19	51.5	22.9	305	1 PPNK_PASMU	O9cnu2 solanum mel
20	51.5	22.9	540	1 YM6M_YEAST	O03263 saccharomyc
21	51.5	22.9	567	1 PCGB_HUMAN	O96149 mus sapien
22	51	22.7	295	1 PPNK_PSEAE	O9h3c0 pseudomonas
23	51	22.7	452	1 TPE_DROME	P48613 drosophila
24	50.5	22.4	186	1 NDKM_MOUSE	O9w84 mus musculu
25	50.5	22.4	205	1 RACC_DICDI	O99q99 dictyostell
26	50.5	22.4	294	1 PPNK_VIBCH	O9ktp8 vibrio chol
27	50.5	22.4	423	1 GRPR_PIG	P34999 sus scrofa
28	50.5	22.4	504	1 CPDI_RAT	P10633 rattus norv
29	50.5	22.4	504	1 CPDI_RAT	P24456 mus musculu
30	50.5	22.4	505	1 CPDB_MOUSE	P23457 mus musculu
31	50	22.2	166	1 STEF_ECOLI	O30919 escherichia
32	50	22.2	236	1 SLR2_RALSO	O30919 ralsstonia s
33	50	22.2	329	1 Y461_CHLTR	O84467 chlamydia t

34	50	22.2	329	1 Y746_CHLMO	O9p122 chlamydia m
35	50	22.2	354	1 FXF1_HUMAN	O12946 homo sapien
36	50	22.2	454	1 AATM_LUPAN	P26563 lupinus ang
37	50	22.2	636	1 P73_HUMAN	O15520 homo sapien
38	50	22.2	693	1 S6A3_BOVIN	P27322 bos taurus
39	50	22.2	1586	1 SN22_HUMAN	P51531 homo sapien
40	49.5	22.0	305	1 PPN1_SYNX3	P74430 synecocyst
41	49.5	22.0	487	1 XYLC_PSEPU	P43503 pseudomonas
42	49.5	22.0	504	1 CPD9_MOUSE	P11714 mus musculu
43	49.5	22.0	897	1 CYRB_HUMAN	P32927 homo sapien
44	49	21.8	277	1 PPNK_PYRAB	O9v081 pyrococcus
45	49	21.8	292	1 PPNK_BUCAT	P57282 buchnera ap

ALIGNMENTS

CAUP_DROME	STANDARD:	PRT:	693 AA.
ID	CAUP_DROME		
AC	P54269: O9VU00:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Homeobox protein caupolican.		
GN	CAUP OR CG10605.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;		
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_Taxid=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96180722; PubMed=8620542;		
RA	Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,		
RA	Ferres-Marco D., Modolell J.;		
RT	*Araucan and caupolican, two members of the novel iroquois complex,		
RT	encode homeoproteins that control proneural and vein-forming genes.;		
RL	Cell 85:95-110(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Berkely;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amantides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abrill J.F., Abpayani A., An H.-J., Andrews-Plankkoc C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Domnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,		
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkllov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,		

```

RA SVtkskasr, Tector C., Turner R. Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarom D.A., Meintock G.M., Weissenhch J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhao M., Zhang G., Zhou L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -I- FUNCTION: CONTROLS PROLIFERATION AND VEIN FORMING GENES. POSITIVE
CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
CC TRANSCRIPTION INITIATION.
CC -I- SUBCELLULAR LOCATION: Nuclear (Probable).
CC MISCELLANEOUS: 'CAUDOLICAN' IS NAMED AFTER THE ARACUANIAN
CC AMERICAN-INDIAN TRIBE, ALSO CALLED MOHAWKS, WHO SHAVED ALL BUT A
CC MEDIAL STRIP OF HAIRS ON THE HEAD.
CC -I- SIMILARITY: BELONGS TO THE TAIE/TRO HOMEBOX FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95178; CA644485.1; -.
DR EMBL; AE003540; AAFA9895.1; -.
DR HSSP; P41778; ID06.
DR FlyBase; FBgn0015919; caup.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR003893; Iroquois-homeo.
DR Pfam; PF00046; homeobox_1.
DR ProDom; PD000010; Homeobox_1.
DR SMART; SM00389; HOX; 1.
DR START; SM00548; IRO; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50077; HOMEBOX_2; 1.
KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW Developmental protein.
FT FT DNAS_BIND 226 288 HOMEBOX (TALE-TYPE).
FT FT DOMAIN 300 303 POLY-ASP.
FT FT DOMAIN 405 418 POLY-GLN.
FT FT DOMAIN 501 516 POLY-GLN.
FT FT DOMAIN 517 528 POLY-HIS.
FT FT DOMAIN 565 572 POLY-SER.
FT FT DOMAIN 613 624 POLY-SER.
FT FT CONFLICT 106 106 C -> R (IN REF. 1).
FT FT CONFLICT 316 316 G -> A (IN REF. 1).
FT FT CONFLICT 678 678 G -> A (IN REF. 1).
SQ SEQUENCE 693 AA; 73667 MW; FEBB161d493f7ECG CRC64;

Query Match 28.0%; Score 63; DB 1; Length 693;
Best Local Similarity 41.2%; Pred. No. 2.6;
Matches 14; Conservative 7; Mismatches 9; Indels 4; Gaps 2;

Oy 7 LCPTGGCGVAFGGCPYPMPQFTQSMRYTLVV 40
Db 141 MALPSGVGVGYGG---PYPSNEQN-PYPSIGV 170

RESULT 2
PPNK_AQUAE STANDARD: PRT: 274 AA.
AC 067055;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23)
DE (Poly(P)/ATP NAD kinase).
PPNK OR AO_909.
NS Aquifex aeolicus.
```

CC Bacteri; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anuj M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus".
RL Nature 392:353-358(1998).
CC -I- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC polyphosphates as a source of phosphorus (by similarity).
CC -I- CATALYTIC ACTIVITY: ATP + NAD(+) -> ADP + NADP(+).
CC -I- COFACTOR: Requires divalent metal ions for activity (by
CC similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000713; AAC07026.1; -
DR InterPro; IPR002504; ATP_NADK.
DR Pfam; PF01513; NAD_kinase; 1.
KW Transferrase; Kinase; NAD; NADP; Complete proteome.
SQ SEQUENCE 274 AA; 31327 MW; 34510FB30EB82C6 CRC64;

Query Match 27.88; Score 62.5; DB 1; Length 274;
Best Local Similarity 35.18; Pred. NO. 1.2;
Matches 13; Conservative 6; Mismatches 13; Indels 5; Gaps 1;

OY 6 L L C L P T G G P V A I F G C G P V P M P Q - - - - F T G S M P P T L 37
Db 166 I L S T P T G S T A Y A L S A G R I P V P E S Q N L L F Y P I C H T L 202

RESULT 3
PPNK_STRCO
-ID PPNK_STRCO STANDARD; PRT: 301 AA.
AC 09S219;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probably inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23)
DE (Poly(P)/ATP NAD kinase).
PPNK OR SC01781 OR SC151.21C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RP SEQUENCE FROM N.A.
RC STRAIN=AD3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kreser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Crocin A., Fraser A., Goble A., Hildgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kisser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rajnowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces


```

RT coelicolin A3(2).";
RL Nature 417:141-147(2002).
CC -1- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC polyphosphate as a source of phosphorus (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + NAD(+) = ADP + NADP(+).
CC -1- COFACTOR: Requires divalent metal ions for activity (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL109848; CAB52845.1; -.
DR InterPro: IPR002504; ATP_NADK.
DR Pfam: PF01513; NAD_kinase; 1.
KW Transferase; Kinase; NAD; NADP; Complete proteome.
SQ SEQUENCE 301 AA; 32081 MW; FCB6F6B01766AEB4 CRC64;

Query Match 27.6%; Score 62; DB 1; Length 301;
Best Local Similarity 47.8%; Pred. No. 1.5;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 6 LCLPTGPGGVAIFGGGPPVMPQ 28
DB 184 VLSTPTGTAFAFAGGPPVWPE 206
: 1 111 1 111 11:
: 1 111 1 111 11:

RESULT 4
ID UP08_ECOLI STANDARD; PRT; 294 AA.
AC P39173; P76233;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page (Spots T26/PR37).
GN YEAD OR B1780.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97251358; PubMed-9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubraman S., Tagami H.,
RA Takada K., Takemoto K., Wada C., Yamamoto Y., Horiiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [3]
RP SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;

```

```

RX MEDLINE-97443975; PubMed-9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE-99420866; PubMed-10493123;
RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;
RT "Enrichment of low abundance proteins of Escherichia coli by
RT hydroxyapatite chromatography.";
RL Electrophoresis 20:2181-2195(1999).
CC -1- SIMILARITY: BELONGS TO THE UPF0010 FAMILY. SIGNIFICANTLY RELATED
CC TO MUTAROTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000273; AAC74850.1; ALT_INIT.
DR EMBL: D90821; BAA15577.1; -.
DR EMBL: D90822; BAA15581.1; -.
DR Ecogen: EC12679; Yead.
DR InterPro: IPR001823; Ald1-epimerase.
DR Pfam: PF01263; Aldose-epim; 1.
KW Complete proteome.
FT ACT_SITE 164 164 POTENTIAL.
SQ SEQUENCE 294 AA; 32666 MW; ADACFEDD46F20957 CRC64;

Query Match 26.4%; Score 59.5; DB 1; Length 294;
Best Local Similarity 38.9%; Pred. No. 3;
Matches 14; Conservative 4; Mismatches 5; Indels 13; Gaps 1;

QY 15 GVAIFGGGPPVMP-----QFTSMPTL 37
DB 70 GVAIRGCVPCWPFGRPAQGLPARGFARNLPWTL 105
: 1111111111 1 ::::1:11
: 1111111111 1 ::::1:11

RESULT 5
ID CHAC_ECOLI STANDARD; PRT; 231 AA.
AC P39163; P77176;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cation transport protein chac.
GN CHAC OR B1218.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NM8191;
RX STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-K12;
RA MEDLINE-97061202; PubMed-8905232;
RA Oshima T., Baba H., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sawai G., Seki Y., Tagami H., Takekoto K., Wada C., Yamamoto Y.,
RA Yano M., Horinouchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA, Res. 3:137-155(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EM3L, L28709; AAA20198.1; ALT_INIT.
CC EM3L, AE000220; AAC74302.1; ALT_INIT.
CC EMBL; D90756; BAA36076.1; ALT_INIT.
CC EMBL; D90757; BAA36086.1; ALT_INIT.
CC PccGene; EG12403; Chac.
CC Transprot; Complete Proteome.
CC TRANSLOCATOR 15 15 A -> V (IN REF. 1).
CC CONFLICT 79 79 C -> F (IN REF. 1).
CC FT CONFLICT 79 79 C -> F (IN REF. 1).
CC FT CONFLICT 79 79 C -> F (IN REF. 1).
CC SO SEQUENCE 231 AA; 25512 MW; 9D708565D53A7C CRC64;

Query Match 26.2%; Score 59; DB 1; Length 231;
Best Local Similarity 47.2%; Pred. No. 2.7;
Matches 17; Conservative 3; Mismatches 12; Indels 4; Gaps 2;

OY 6 LCLPTGPGVAIFGCGPVPWP---OFTOSMPYTLV 38
DB 38 LACRPDEGP-VWIFGYGSLMMNPALETETESCTGLV 72

RESULT 6
PNPK_MYCTU STANDARD; PRT; 307 AA.
AC 033196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23) (Poly(P)/ATP NAD
DE kinase)
GN PNPK OR RV1695 OR MT1734 OR MTC1125.17.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN-H37RV;
RA MEDLINE-20462915; PubMed-11006082;
RA Kawai S., Mori S., Mukai T., Suzuki S., Yamada T., Hashimoto W.,
RA Murata K.;
RT "Inorganic polyphosphate / ATP-NAD kinase of Micrococcus flavus and
RT Mycobacterium tuberculosis H37RV.";
RL Biochem. Biophys. Res. Commun. 276:57-63(2000).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekla F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

```

```

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC polyphosphate as a source of phosphorus.
CC -1- CATALYTIC ACTIVITY: ATP + NAD(+) -> ADP + NADP(+).
CC -1- COFACTOR: Requires divalent metal ions for activity.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB044336; BAB21478.1; -
CC EMBL; 298268; CAB10952.1; -
CC EMBL; AE007035; AAK66003.1; -
CC TIGR; MT1734;
DR TuberculList; RV1695; -
DR InterPro; IPR002504; ATP_NADK.
DR Pfam; PF01513; NAD_kinase; 1.
DR Transferrase; Kinase; NADP: Complete proteome.
SO SEQUENCE 307 AA; 32903 MW; 68817BE570B6645B CRC64;

Query Match 26.2%; Score 59; DB 1; Length 307;
Best Local Similarity 50.0%; Pred. No. 3.6;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 6 LCLPTGPGVAIFGCGPVPWP 27
DB 192 LVSTPTGTAFAFSAGGVLMP 213

RESULT 7
PNPK_MYCLE STANDARD; PRT; 311 AA.
AC 049897; 005675;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23) (Poly(P)/ATP NAD
DE kinase)
GN PNPK OR ML1359 OR MLC1351.13C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-TN;
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-TN;
RA MEDLINE-21128732; PubMed-11234002;

```

RA Cole S.T., Elismeyer K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holtroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -i- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
 CC ATP and other nucleoside triphosphates as well as inorganic
 CC polyphosphate as a source of phosphorus (By similarity).
 CC -i- CATALYTIC ACTIVITY: ATP + NAD(+) = ADP + NADP(+).
 CC -i- COFACTOR: Requires divalent metal ions for activity (By
 CC similarity).
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -i- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U00021; AAA50923.1; -
 CC DR EMBL: Z95117; CAB08286.1; -
 CC DR EMBL: AL583921; CAC31740.1; -
 CC DR Lepidoma; ML1359; -
 CC DR InterPro: IPR002504; ATP_NADK.
 CC DR Pfam: PF01513; NAD_kinase; 1.
 CC KW Transferrase; Kinase; NAD; NADP; Complete proteome.
 CC SQ SEQUENCE 311 AA; 3328 MW; 0AD22150C62D3B CRC64;
 QY
 Query Match 26.2%; Score 59; DB 1; Length 311;
 Best Local Similarity 50.0%; Pred. No. 3.7;
 Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 Db 6 LCLPTGGPVAIFGGGPVWP 27
 196 LVSTPTGGTAFAFSGGAPVLP 217
 ID YXCE_BACSU STANDARD; PRT; 216 AA.
 YXCE_BACSU
 AC P46335;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yxce.
 GN YXCE OR SS92DR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RX MEDLINE=96093926; PubMed=7584049;
 RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
 RL genome between the gnt and lol operons.";
 RL DNA Res. 2:61-69(1995).
 RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Belterio M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.M., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Gallion N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Goldightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwolik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takauchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Wetzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumslein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -i- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AB005554; BAA21606.1; -
 CC DR EMBL: Z99124; CAB16015.1; -
 CC DR Subtilisin; BG11362; yxce.
 CC KW Hypothetical protein; Transmembrane; Complete proteome.
 CC FT TRANSMEM 5
 CC FT TRANSMEM 137 157
 CC FT TRANSMEM 183 203
 CC FT DOMAIN 34 46
 CC FT DOMAIN 57 67
 CC FT DOMAIN 76 81
 CC FT DOMAIN 91 95
 CC FT POLY-SER.
 CC FT POLY-SER.
 CC FT POLY-SER.
 CC SQ SEQUENCE 216 AA; 23074 MW; 1B850B1D9EDCA42 CRC64;
 QY
 Query Match 25.8%; Score 58; DB 1; Length 216;
 Best Local Similarity 47.1%; Pred. No. 3.4;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 Db 10 PTGGPVAIFGGGPVWP 26
 169 PVGGPVTYGCASPIAW 185
 ID PGGB_MOUSE STANDARD; PRT; 568 AA.
 PGGB_MOUSE
 AC Q8VHQ3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
 GN inhibited membrane-associated protein) (CAAX box protein TIMAP).
 GN PP1R16B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RA	Stalmans W., Bollen M.
RT	"Molecular cloning of NIPP-1, a nuclear inhibitor of protein phosphatase-1, reveals homology with polypeptides involved in RNA processing."
RT	J. Biol. Chem. 270:28068-28074(1995).
RN	121
RP	PP-1 BINDING DOMAIN, AND SYNTHESIS OF PEPTIDES IN THE 141-230 REGION.
RX	MEDLINE=99253961; Pubmed=10318819.
RA	Buellens M., Van Eynde A., Valsteke V., Connor J., Shenolikar S., Stalmans W., Bollen M.
RT	"Molecular determinants of nuclear protein phosphatase-1 regulation by NIPP-1".
RT	J. Biol. Chem. 274:14053-14061(1999).
CC	-1 FUNCTION: INHIBITOR SUBUNIT OF THE MAJOR NUCLEAR PROTEIN PHOSPHATASE-1 (PP-1). IT HAS RNA BINDING ACTIVITY BUT DOES NOT CLEAVE RNA AND MAY TARGET PP-1 TO RNA-ASSOCIATED SUBSTRATES.
CC	-1 SUBCELLULAR LOCATION: Nuclear.
CC	-1 ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B: ARE PRODUCED BY ALTERNATE SPLICING.
CC	-1 TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED. HIGHEST EXPRESSION FOUND IN HEART AND SKELETAL MUSCLE, FOLLOWED BY BRAIN, PLACENTA, LUNG, LIVER AND PANCREAS. LESS ABUNDANT IN KIDNEY.
CC	-1 DOMAIN: HAS A BASIC N- AND C-TERMINAL AND AN ACIDIC CENTRAL DOMAIN.
CC	-1 PTM: THE N-TERMINAL IS BLOCKED.
CC	-1 PTM: DOWN-REGULATED BY PHOSPHORYLATION (PROBABLE).
CC	-1 SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL: Z50748; CAA90625.1; .
DR	InterPro: IPR000253; FHA_domain.
DR	Pfam: PF00498; FHA: 1.
DR	SMART: SM00240; FHA: 1.
DR	PROSITE: PSS0006; FHA_DOMAIN: 1.
KW	Protein phosphatase inhibitor; RNA-binding; Alternative splicing;
KM	Phosphorylation.
FT	DOMAIN 2 5 POLY-ALA.
FT	DOMAIN 49 101 FHA.
FT	DOMAIN 191 200 INVOLVED IN PP-1 INHIBITION.
FT	DOMAIN 200 203 INVOLVED IN PP-1 BINDING.
FT	MOD_RES 199 199 PHOSPHORYLATION (PROBABLE).
FT	DOMAIN 331 334 POLY-LYS.
FT	VARSPLIC 1 142 MISSING (IN ISOFORM B).
SQ	SEQUENCE 351 AA; 38521 MW; 7CDSF6E162A66210 CRC64;
Query Match	25.3% Score 57; DB 1; Length 351;
Best Local Similarity	40.0%; Pred. No. 7.3;
Matches 12; Conservative	4; Mismatches 14; Indels 0; Gaps 0;
DQ	10 PTGSGPVAlfGGCPVWPOTQSMPRLV 39 : : : : DQ 278 PHGIHGTAfLGCLPMYPNLAPVDVLPRV 307
RESULT 11	
ID	PPR8_HUMAN STANDARD: PRT: 351 AA.
AC	Q12972; Q9UBZ0; Q9UBH1.
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Nuclear inhibitor of protein phosphatase-1 (NIPP-1) (Protein phosphatase 1, regulatory inhibitor subunit 8) [includes: Activator of RNA decay (PC 3.1.4.-) (ARD-1)].
DE	PEP18 OR NIP1 OR ARD1.
GN	Homo sapiens (Human).

RC STRAIN-ATCC 33303 / B10;
 RX MEDLINE-95362677; PubMed-7635831;
 RA Pollich M., Kling G.;
 RT "Identification and sequence analysis of genes involved in late steps
 in cobalamin (vitamin B12) synthesis in Rhodobacter capsulatus.";
 RL J. Bacteriol. 177:4481-4487(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SB1003 / St Louis;
 RX MEDLINE-97404404; PubMed-9256491;
 RA Vlack C., Paces V., Maltsev N., Paces J., Haselkorn R., Feinstein M.;
 RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
 capsulatus SB1003.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF COBRYIC ACID TO
 COBANAMIDE. ADDITION OF AMINOPROPANOL ON THE F CARBOXYLIC
 GROUP (BY SIMILARITY).
 CC -1- PATHWAY: Cobalamin biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE COBD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z46611; CAA86581.1;
 DR EMBL: AF010496; AAC16177.1;
 DR InterPro: IPR004485; Cblb.
 DR Pfam: PF03186; COBD_Cblb; 1.
 DR TIGRfams: TIGR00380; cblb; 1.
 DR Ccobalamin biosynthesis; Transmembrane.
 KW TRANSMEM 3
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 78 98 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 290 310 POTENTIAL.
 FT CONFLICT 35 36 GA -> AR (IN REF. 1).
 SQ SEQUENCE 314 AA; 33407 MW; DE5F592A91DB46D8 CRC64;
 QY 7 LCIPGCGPVAFRGGVPMP 27
 Db 76 LWLPAGWPGVLI--GGILAMP 94
 RESULT 13
 CRAR_HUMAN STANDARD: PRT: 699 AA.
 ID F48740; O95570; O9UF09;
 AC F48740; O95570; O9UF09;
 DT 1-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement-activating component of Ra-reactive factor precursor
 (EC 3.4.21.-) (Ra-reactive factor serine protease p100) (RARF)
 DE (Mannan-binding lectin serine protease 1) (Mannose-binding protein
 associated serine protease) (MASP-1).
 DE GN HAPSI OR CRARF OR CRARF1 OR PRSS5.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE-94059062; PubMed-8240317;
 RA Takada F., Takayama Y., Hattuse H., Kawakami M.;

RT "A new member of the C1s family of complement proteins found in a
 bactericidal factor, Ra-reactive factor, in human serum.";
 RT Biochem. Biophys. Res. Commun. 196:1003-1009(1993).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Fetal liver;
 RX MEDLINE-94289349; PubMed-8018603;
 RA Sato T., Endo Y., Matsushita M., Fujita T.;
 RT "Molecular characterization of a novel serine protease involved in
 activation of the complement system by mannose-binding protein.";
 RL Int. Immunol. 6:665-669(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE-97079701; PubMed-8921412;
 RA Endo Y., Sato T., Matsushita M., Fujita T.;
 RT "Exon structure of the gene encoding the human mannose-binding
 protein-associated serine protease light chain: comparison with
 complement C1r and C1s genes.";
 RL Int. Immunol. 8:1355-1358(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94023590; PubMed-10475605;
 RA Takayama Y., Takada F., Nowatari M., Kawakami M., Matsuura N.;
 RT "Gene structure of the p100 serine-protease component of the human Ra-
 reactive factor.";
 RL Mol. Immunol. 36:505-514(1999).
 CC -1- FUNCTION: COMPONENT OF THE BACTERICIDAL RA-REACTIVE FACTOR RARF
 WHICH SPECIFICALLY BINDS TO RA AND R2 POLYSACCHARIDES EXPRESSED BY
 CERTAIN ENTEROBACTERIA. IT TRIGGERS THE ACTIVATION OF COMPLEMENT
 CASCADE BY ACTIVATING THE C4 AND C2 COMPONENTS. IT ACTIVATES THE
 C4 COMPONENT BY CLEAVING THE ALPHA-CHAIN OF C4.
 CC -1- SUBUNIT: RARF CONSISTS OF A COMPLEMENT-ACTIVATING COMPONENT
 (CRARF) AND A POLYSACCHARIDE-BINDING (MANNOSE-BINDING) COMPONENT.
 CC CRARF IS AN HETERODIMER OF A HEAVY (P70) AND A LIGHT CHAIN (P29)
 CC LINKED BY A DISULFIDE BOND.
 CC -1- DOMAIN: CRARF HAS A MODULE ORGANIZATION SIMILAR TO C1R AND C1S.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D17525; BAA04477.1;
 DR EMBL: D28593; BAA05928.1;
 DR EMBL: D61695; BAA34864.1;
 DR EMBL: AB010822; BAA34864.1; JOINED.
 DR EMBL: AB010813; BAA34864.1; JOINED.
 DR EMBL: AB010814; BAA34864.1; JOINED.
 DR EMBL: AB010815; BAA34864.1; JOINED.
 DR EMBL: AB010816; BAA34864.1; JOINED.
 DR EMBL: AB010817; BAA34864.1; JOINED.
 DR EMBL: AB010818; BAA34864.1; JOINED.
 DR EMBL: AB010819; BAA34864.1; JOINED.
 DR EMBL: AB010820; BAA34864.1; JOINED.
 DR EMBL: AB010821; BAA34864.1; JOINED.
 DR EMBL: D61690; BAA34864.1; JOINED.
 DR EMBL: D61691; BAA34864.1; JOINED.
 DR EMBL: D61692; BAA34864.1; JOINED.
 DR EMBL: D61693; BAA34864.1; JOINED.
 DR EMBL: D61694; BAA34864.1; JOINED.
 DR EMBL: AB007617; BAA89206.1;
 DR EMBL: AB007602; BAA89206.1; JOINED.
 DR EMBL: AB007603; BAA89206.1; JOINED.
 DR EMBL: AB007604; BAA89206.1; JOINED.
 DR EMBL: AB007605; BAA89206.1; JOINED.

DR EMBL: AB007606; BAA89206.1; JOINED.
 DR EMBL: AB007607; BAA89206.1; JOINED.
 DR EMBL: AB007608; BAA89206.1; JOINED.
 DR EMBL: AB007609; BAA89206.1; JOINED.
 DR EMBL: AB007610; BAA89206.1; JOINED.
 DR EMBL: AB007611; BAA89206.1; JOINED.
 DR EMBL: AB007612; BAA89206.1; JOINED.
 DR EMBL: AB007613; BAA89206.1; JOINED.
 DR EMBL: AB007614; BAA89206.1; JOINED.
 DR EMBL: AB007615; BAA89206.1; JOINED.
 DR EMBL: AB007616; BAA89206.1; JOINED.
 DR HSP: P00736; IAP0.
 DR MEROPS; S01.198; -.
 DR Genew; HGNC:6901; MASP1.
 DR MIM: 600521; -.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000859; CUB_domain.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000361; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001254; Ser-protease_Try.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR Pfam: PF00084; sushi; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PF00431; CUB; 2.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00032; CCP; 2.
 DR SMART: SM00042; CUB; 2.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00179; EGF_Ca; 1.
 DR SMART: SM00020; TRYP_SPE; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01186; EGF; 2.
 DR PROSITE: PS01187; EGF_Ca; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydroxylase; Complement pathway; Serine protease; Protease;
 KM Glycoprotein; Sushi; Repeat; Signal; EGF-like domain; Hydroxylation.
 FT SIGNAL 1 19
 FT CHAIN 20 699
 FT FT
 FT CHAIN 20 448
 FT CHAIN 449 699
 FT DOMAIN 20 138
 FT DOMAIN 139 182
 FT DOMAIN 185 297
 FT DOMAIN 300 363
 FT DOMAIN 366 433
 FT DOMAIN 449 699
 FT ACT_SITE 490 490
 FT ACT_SITE 552 552
 FT ACT_SITE 646 646
 FT MOD_RES 159 159
 FT MOD_RES 73 91
 FT DISULFID 143 157
 FT DISULFID 153 166
 FT DISULFID 168 181
 FT DISULFID 185 212
 FT DISULFID 242 260
 FT DISULFID 301 349
 FT DISULFID 329 362
 FT DISULFID 367 414
 FT DISULFID 397 432
 FT DISULFID 436 572
 FT DISULFID 614 631
 FT DISULFID 642 672
 FT CARBOHYD 49 49
 FT CARBOHYD 178 178
 FT CARBOHYD 385 385
 FT CARBOHYD 407 407
 FT CONFLICT 235 235

Q -> E (IN REF. 1 AND 4).

FT CONFLICT 285 285 A -> G (IN REF. 1).
 FT CONFLICT 499 499 K -> E (IN REF. 3).
 FT CONFLICT 499 499 K -> E (IN REF. 2).
 FT CONFLICT 527 527 D -> A (IN REF. 3).
 FT CONFLICT 543 543 Q -> K (IN REF. 1).
 FT CONFLICT 352 352 D -> V (IN REF. 3).
 FT CONFLICT 643 643 A -> S (IN REF. 1).
 FT SEQUENCE 699 AA: 79258 MW: 74AE180C6E0AF0D5 CRC64;
 SQ
 Query Match 23.6%; Score 53; DB 1; Length 699;
 Best Local Similarity 36.2%; Pred. No. 44;
 Matches 17; Conservative 6; Mismatches 14; Indels 10; Gaps 3;
 QY 1 VANRFL--LCLEPRG---GPGVAIFGGGYPWPQFTQSMPTLVVVK 41
 DB 563 VNAFVFWPCLPEGPQEGAMVYSGWG---KQFLRPREETIMLE 605
 RESULT 14
 ID IRXH_CAEEL STANDARD; PRT; 356 AA.
 AC 093348;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative Iroquois-class homeobox protein C36F7.1.
 GN C36F7.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermatidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=Bristol N2;
 RA Lightning J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RP REVISIONS.
 RA Jones S.J.M.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE TALE/IRO HOMEBOX FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z81045; CAB02816.1; -.
 DR HSP: P40424; 1B72.
 DR TRANSFAC: T04278; -.
 DR Wormpep: C36F7.1; CE17526.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR003893; Iroquois_homeo.
 DR Pfam: PF00046; homeobox; 1.
 DR Prodom: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR SMART: SM00548; IRO; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 KW Hypothetical protein; Homeobox; DNA-binding; Nuclear protein.
 FT DOMAIN 54 63 POLY-ALA.
 FT DNA_BIND 115 177 HOMEBOX (TALE-TYPE).
 FT DOMAIN 192 201 POLY-ASP.
 SQ SEQUENCE 356 AA: 38689 MW: 74AE180C6E0AF0D5 CRC64;
 Query Match 23.3%; Score 52.5; DB 1; Length 356;
 Best Local Similarity 32.7%; Pred. No. 26;
 Matches 16; Conservative 2; Mismatches 12; Indels 19; Gaps 2;

QY 12 GCGPVAIF-----GCGPVPMPQFTQSM-----PYTLVYVK 41
 Db 70 GCGPVPMPHPCIPADLPKPEMLLGGCGPMPMFSDAHLVHPYGLDGIK 118

RESULT 15
 PAC1_MOUSE STANDARD; PRT; 441 AA.

AC 061644;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Protein kinase C and casein kinase substrate in neurons protein 1.
 GN PACSIN1 OR PACSIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=98417447; PubMed=9746365;
 RA Plomann M., Lange R., Vopper G., Cramer H., Heinlein U.A.O.,
 RA Scheif S., Baldwin S.A., Lettges M., Cramer M., Paulsson M.,
 RA Bartheis D.;
 RT "PACSIN, a brain protein that is upregulated upon differentiation into
 RT neuronal cells.";
 RL Eur. J. Biochem. 256:201-211(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Eye;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION
 RX PubMed=11082044;
 RA Modregger J., Rilter B., Witter B., Paulsson M., Plomann M.;
 RT "All three PACSIN isoforms bind to endocytic proteins and inhibit
 RT endocytosis.";
 RL J. Cell Sci. 113:4511-4521(2000).
 CC -1- FUNCTION: May play a role in vesicle formation and transport.
 CC -1- SUBUNIT: Homo- and hetero-aggregates with other PACSINs. Binds
 CC dynamin I, synaptobrevin, synapsin I and the neural Wiskott-Aldrich
 CC syndrome protein (N-WASP) (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Colocalizes with dynamin I at
 CC vesicular structures in the cell body and neurites (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in CNS neurons, especially in
 CC the pyramidal cells of the hippocampus, Purkinje cells of the
 CC cerebellum and large neurons of the cortex and brain stem.
 CC -1- DEVELOPMENTAL STAGE: Expression is seen at embryonic day 17 and is
 CC upregulated developmentally with a correlation to neuronal
 CC differentiation.
 CC -1- PTM: Phosphorylated by casein kinase 2 (CK2) and protein kinase C
 CC (PKC) (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PACSIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 FCH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC
 CC EMBL: X85124; CAA59437.1;
 DR EMBL: BC014698; AAH14698.1;
 DR HSSP: G60631; IGBO.
 DR MGD: MGI:1345181; PACSIN1.
 DR InterPro: IPR001060; Cdc15_Fes_CTP4.
 DR InterPro: IPR001452; SH3.

DR Pfam: PF00611; FCH; 1.
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR SMART: SM00055; FCH; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50133; FCH; 1.
 DR PROSITE: PS50002; SH3; 1.
 KW Endocytosis; SH3 domain; Coiled coil; Phosphorylation.
 FT DOMAIN 10 73 FCH.
 FT DOMAIN 382 441 SH3.
 FT DOMAIN 144 165 COILED COIL (POTENTIAL).
 FT DOMAIN 183 217 COILED COIL (POTENTIAL).
 SQ SEQUENCE 441 AA; 50575 MW; 21BBB339A14A41F9 CRC64;

Query Match 23.3%; Score 52.5; DB 1; Length 441;
 Best Local Similarity 37.9%; Pred. NO. 32;
 Matches 11; Conservative 5; Mismatches 4; Indels 9; Gaps 2;

QY 11 TCGPVAIFGCGPVPMPQFTQ---SMPT 36
 Db 287 TSGPGM-----PMNMPQFEEMNDPLPT 309

Search completed: May 20, 2003, 16:35:47
 Job time : 12.0385 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:19:09 ; Search time 50.6868 Seconds
(without alignments)
166.669 Million cell updates/sec

Title: US-09-869-155-19
Perfect score: 225
Sequence: 1 VANRFLCLPTGPGVAIFG.....GVPMPQFTQSMPTLVVK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP. archaea: *
2: SP. bacteria: *
3: SP. fungi: *
4: SP. human: *
5: SP. invertebrate: *
6: SP. mammal: *
7: SP. mhc: *
8: SP. organelle: *
9: SP. phage: *
10: SP. plant: *
11: SP. rodent: *
12: SP. virus: *
13: SP. vertebrate: *
14: SP. unclassified: *
15: SP. virus: *
16: SP. bacteriophage: *
17: SP. archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	49.3	419	10	08S1V0
2	96	42.7	424	10	08S1V1
3	79	35.1	422	10	08S1V3
4	77.5	34.4	424	10	08S1V5
5	71	31.6	402	10	08S1U9
6	68	30.2	492	16	08XT86
7	67.5	30.0	434	10	08S1U4
8	61	27.1	362	2	08V0U9
9	61	27.1	446	16	09A7W1
10	59.5	26.4	301	16	08XD06
11	59.5	26.4	433	10	09ZVS4
12	59.5	26.4	534	17	097XY4
13	59	26.2	434	10	09ZVS5
14	57.5	25.6	568	11	08VH03
15	57	25.3	263	4	09NV23
16	57	25.3	351	11	08R3G1

17	56.5	25.1	283	16	08RAC3	08rac3 thermomane
18	56.5	25.1	1054	5	016331	016331 caenorhabd
19	56	24.9	291	16	08REK9	08REK9 yersinia pe
20	56	24.9	305	2	053129	053129 rhodobacter
21	56	24.9	650	8	09XPH4	09XPH4 beta vulgar
22	55.5	24.7	750	11	09CW61	09CW61 mus musculu
23	55.5	24.7	1050	5	016436	016436 caenorhabd
24	55	24.4	365	12	08ORT0	08ORT0 chimpanzee
25	55	24.4	629	2	007639	007639 streptomyce
26	54.5	24.2	193	11	099ML1	099ML1 mus musculu
27	54.5	24.2	328	5	095YD3	095YD3 caenorhabd
28	54	24.0	195	16	08XSR5	08XSR5 raiistonla s
29	54	24.0	232	4	09H2W2	09H2W2 homo sapien
30	54	24.0	288	16	09PB74	09PB74 xylella fas
31	54	24.0	331	2	087941	087941 thauera aro
32	54	24.0	546	4	096M29	096M29 homo sapien
33	54	24.0	559	4	09H5P1	09H5P1 homo sapien
34	54	24.0	621	10	09LGX7	09LGX7 oryza sativ
35	54	24.0	1302	4	09C092	09C092 homo sapien
36	54	24.0	1980	5	09NDY5	09NDY5 leishmania
37	54	24.0	2552	2	09XCF2	09XCF2 mycobacteri
38	53.5	23.8	325	16	09FD03	09FD03 xylella fas
39	53.5	23.8	463	10	039496	039496 cylindrothe
40	53.5	23.8	671	5	09VB83	09VB83 drosophila
41	53.5	23.8	684	5	024035	024035 drosophila
42	53.5	23.8	768	16	086766	086766 streptomyce
43	53.5	23.8	831	5	08TA47	08TA47 drosophila
44	53	23.6	293	11	09D6L4	09D6L4 mus musculu
45	53	23.6	294	2	09R6G6	09R6G6 agrobacteri

ALIGNMENTS

RESULT 1	
08S1V0	PRELIMINARY; PRT; 419 AA.
AC 08S1V0	
DT 01-JUN-2002 (TREMBLrel. 21, Created)	
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE Putative dermal glycoprotein.	
GN P0504E02.9.	
OS Oryza sativa (japonica cultivar-group).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC Ehrhartoideae; Oryzae; Oryza.	
OX NCBI_TaxID=39947;	
OX [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=CV. NIPPONBARE;	
RA Sasaki T., Matsumoto T., Yamamoto K.;	
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC	
RT clone:P0504E02."	
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL: AP003269; BAB89708.1; -	
SO SEQUENCE	419 AA; 43786 MW; 792B55FAAF3F3A8 CRC64;
Query Match	49.3%; Score 111; DB 10; Length 419;
Best Local Similarity	55.3%; Pred. No. 7.7e-06;
Matches	26; Conservative 4; Mismatches 11; Indels 6; Gaps 2;
QY 1 VANRFLCLPT--GGPGVAIFGGGPV---PMPQFTQSMPTLVVK 41	
DB 195 VAGRFLLCLPRLGYGGVAIFGGGPVYEGSLPDTTLDYPLVAK 241	
RESULT 2	
ID 08S1V1	PRELIMINARY; PRT; 424 AA.
AC 08S1V1	
DT 01-JUN-2002 (TREMBLrel. 21, Created)	
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)	


```

RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Guiguer P., Thibault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.,
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646077; CAD17380.1;
DR InterPro: IPR002086; Aldehyde_dehydr.
DR Pfam: PF00171; aldedh; 1.
KM Oxidoreductase; plasmid; Complete proteome.
SQ SEQUENCE 492 AA; 52058 MW; F795296E3EF7FEC6 CRC64;

Query Match
Best Local Similarity 30.2%; Score 68; DB 16; Length 492;
Matches 16; Conservative 4; Mismatches 9; Indels 6; Gaps 2;

QY 10 PTGGPGVAIFG---GGPWPMPQFTQSMPTLVYVK 41
DB 453 PFGGPGIAGNGTSMGPGADWEVYTO---WRWVTVK 484

RESULT 7
ID 08S1U4 PRELIMINARY; PRT; 434 AA.
AC 08S1U4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE P0504E02.16 protein.
GN P0504E02.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatloideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0504E02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003269; BAB89714.1;
SQ SEQUENCE 434 AA; 46236 MW; 87C36211DD52FF72 CRC64;

Query Match
Best Local Similarity 30.0%; Score 67.5; DB 10; Length 434;
Matches 17; Conservative 4; Mismatches 9; Indels 11; Gaps 2;

QY 3 NRFLCLPTGGPGVAIFGGPVP-----PWPQFTQSMPT 36
DB 200 NKFALCL---PGFAFGDPVYIGTESLGIVNTESLPYT 236

RESULT 8
ID 08VUL9 PRELIMINARY; PRT; 362 AA.
AC 08VUL9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Poly(P)/ATP-NAD kinase.
GN MNK.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 3242;
RA Kawai S., Murata K.;
RT "Direct phosphorylation of NADH by NAD kinase.";
```

```

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB070351; BAB84189.1;
DR InterPro: IPR002504; ATP_NADK.
DR Pfam: PF01513; NAD_kinase; 1.
KM Kinase.
SQ SEQUENCE 362 AA; 39114 MW; 106EC546E58F47E1 CRC64;

Query Match
Best Local Similarity 27.1%; Score 61; DB 2; Length 362;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 LLCLPTGGPGVAIFGGPVPMP 27
DB 188 VLAPPTGTSTAYAFSGCPVWVP 209

RESULT 9
ID 09A7M1 PRELIMINARY; PRT; 446 AA.
AC 09A7M1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Succinylarginine dihydrolase.
GN CC1608.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ufferbach T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005836; AAK23587.1;
DR TIGR: CC1608;
KM Hydrolase; Complete proteome.
SQ SEQUENCE 446 AA; 48314 MW; 406C09EDC6AF7B9C CRC64;

Query Match
Best Local Similarity 27.1%; Score 61; DB 16; Length 446;
Matches 13; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

QY 2 ANRFLCLPTGGPGVAIFGGPVPMPQFTQSMPT 34
DB 178 ANHVRLCAEHGPGVNLFWMGREAWSHWDGRFP 210

RESULT 10
ID 08XD06 PRELIMINARY; PRT; 301 AA.
AC 08XD06;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Orf, hypothetical protein.
GN YEAD OR 22820 OR ECS2489.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
```

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RT Nature 409:529-533(2001).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE005401; AAG56769.1; -;
 DR EMBL: AP002558; BAB35912.1; -;
 DR InterPro: IPR001823; A1d1.epimerase.
 DR Pfam: PF01263; Aldose-epim; 1.
 KW Complete proteome.
 SQ SEQUENCE 301 AA; 33590 MW; 7F56DA678B6DE326 CRC64;
 Query Match 26.4%; Score 59.5; DB 16; Length 301;
 Best Local Similarity 38.9%; Pred. No. 14;
 Matches 14; Conservative 4; Mismatches 5; Indels 13; Gaps 1;
 QY 15 GVAIFGGPVPWP-----OFTGSMPTL 37
 Db 77 GVAIRGVPVCMWFGPAAGGLPAHGFARMLPTL 112
 ID 092VSA PRELIMINARY; PRT; 433 AA.
 AC 092VSA;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F15K9.17 protein (Hypothetical 45.7 kDa protein) (Putative
 DE extracellular dermal glycoprotein EDGP).
 GN F15K9.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Xysoctskala V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
 RA Xysoctskala J., Luos C., Araujo R., Buehler E., Conway A.B.,
 RA Jester K., Feng J., Kim C., Li Y., Shin P., Sun H., Davis R.W.,
 RA Ecker J.R., Federispiel N.A., Theologos A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence.";
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologos A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologos A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RT SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Theologos A.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Theologos A., Ecker J., Davis R.W.,
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Natusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologos A.;
 RT "Full Length cDNA of gene F15K9.17 (GI3850579).";
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RL [17]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Natusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologos A.;
 RT "Full Length cDNA of gene F15K9.17 (GI3850579).";
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RL [18]
 DR EMBL: AC005278; AAC72119.1; -;
 DR EMBL: AF325092; AAK17160.1; -;
 DR EMBL: AY059098; AAL15204.1; -;
 DR EMBL: AY035026; AAK59531.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 433 AA; 45717 MW; 7214FC4B8BA72962 CRC64;
 Query Match 26.4%; Score 59.5; DB 10; Length 433;
 Best Local Similarity 60.0%; Pred. No. 19;
 Matches 12; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 4 RELCLPTGPGVAFPGGP 23
 Db 199 KFAVCL-TSGKGVAFPGGP 217
 ID 097XY4 PRELIMINARY; PRT; 534 AA.
 AC 097XY4;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Nitric oxide reductase, subunit I (cytochrome B) (norB-1).
 GN NORB-1 OR SSO1571.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OC NCBI_TaxID=2287;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE-21332296; PubMed-11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doollittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RL EMBL: AE006771; AAK41784.1; -;
 KW Complete proteome.

SEQ SEQUENCE 534 AA: 58445 MW: 4E9C344AF3984962 CRC64;

Query Match 26.4%: Score 59.5; DB 17; Length 534;

Best Local Similarity 38.7%: Pred. No. 24; Mismatches 12; Conservative 7; Mismatches 9; Indels 3; Gaps 2;

DB 13 GPGCV-AIFGGGPV--PMPQFTQSMPTLVV 40

DB 240 GSGIGAFISGLPIAPWPNTEDQFLWIMI 270

RESULT 13

09ZVS5 PRELIMINARY; PRT: 434 AA.

ID 09ZVS5: 09ZVS5:

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE F15K9.16 protein (Putative extracellular dermal glycoprotein precursor).

GN F15K9.16.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Vysotskaya V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,

RA Kremenetskaya I., Iurov J., Araujo R., Buehler E., Conway A.B.,

RA Dwyer K., Feng C., Kim C., Li Y., Shinn P., Sun H., Davis R.W.,

RA Ecker J.R., Federspiel N.A., Theologis A.;

RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence."

RL Submitted (Jul-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Theologis A.;

RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Theologis A.;

RL Submitted (Nov-1998) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Elyu P., Lee J.M.,

RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,

RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,

RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;

RT "Full length cDNA of gene F15K9.16 (GI:3850580)."

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC005278; AAC72120.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

DR EMBL: AF332411; AAG48774.1;

ID 08VHQ3 PRELIMINARY; PRT: 568 AA.

AC 08VHQ3:

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Protein phosphatase 1 regulatory subunit 16B.

GN PP1R16B.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=CEREBELLUM;

RA Northcutt G.M., Kurschner C.;

RT "A novel mouse protein phosphatase 1 targeting subunit is expressed in three distinct regions of the brain."

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF423761; AAL62093.1;

DR InterPro: IPR002110; ANK.

DR InterPro: IPR002120; Prenyl_site.

DR Pfam: PF00023; ANK_5.

DR PRINTS: PR01415; ANKYRIN.

DR SMART: SM00248; ANK_5.

DR PROSITE: PS50088; ANK_REPEAT; 4.

DR PROSITE: PS00297; ANK_REPEAT_REGION; 1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:35:02 ; Search time 18.9211 seconds
(without alignments)
63.750 Million cell updates/sec

Title: US-09-869-155-19
Perfect score: 225
Sequence: 1 VANFLLCLPTGPGVAIFG.....GPVWPQFTQSPYTLVVK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2-6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2-6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2-6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2-6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2-6/ptodata/1/1aa/PCYUS-COMB.pep.*
6: /cgn2-6/ptodata/1/1aa/Backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	24.2	504	US-09-126-420A-18	Sequence 18, Appl
2	53	23.6	251	US-08-944-483-47	Sequence 47, Appl
3	53	23.6	304	US-09-088-651-2	Sequence 2, Appl
4	52.5	23.3	441	US-08-630-915A-34	Sequence 34, Appl
5	52.5	23.3	635	US-09-081-975-3	Sequence 3, Appl
6	51	22.7	45	US-08-881-450A-16	Sequence 16, Appl
7	51	22.7	452	US-08-317-880-2	Sequence 2, Appl
8	51	22.7	452	US-08-782-396-2	Sequence 2, Appl
9	50.5	22.4	504	US-09-126-420A-23	Sequence 23, Appl
10	50	22.2	218	US-09-081-975-7	Sequence 7, Appl
11	50	22.2	281	US-08-487-748A-9	Sequence 9, Appl
12	50	22.2	281	US-08-487-748A-10	Sequence 10, Appl
13	50	22.2	281	US-08-480-070C-10	Sequence 10, Appl
14	50	22.2	281	US-08-829-525-10	Sequence 10, Appl
15	50	22.2	281	US-08-609-583A-10	Sequence 10, Appl
16	50	22.2	281	US-08-937-399-10	Sequence 10, Appl
17	50	22.2	281	US-09-310-367-10	Sequence 10, Appl
18	50	22.2	281	US-09-032-337-10	Sequence 10, Appl
19	49.5	22.0	433	US-09-046-158A-2	Sequence 2, Appl
20	49.5	22.0	897	US-07-960-389-2	Sequence 2, Appl
21	49	21.8	1266	US-08-506-296B-4	Sequence 4, Appl
22	48.5	21.6	543	US-09-535-008-63	Sequence 63, Appl
23	48.5	21.6	577	US-08-996-138-61	Sequence 61, Appl
24	48.5	21.6	625	US-08-996-138-15	Sequence 15, Appl
25	48.5	21.6	625	US-08-995-659-15	Sequence 15, Appl
26	48.5	21.6	625	US-09-215-649A-15	Sequence 15, Appl
27	48.5	21.6	625	US-09-577-780-15	Sequence 15, Appl

28	48.5	21.6	1088	US-09-082-059-2	Sequence 2, Appl
29	48.5	21.6	1647	US-09-535-008-67	Sequence 67, Appl
30	48.5	21.6	1647	US-09-535-008-2	Sequence 2, Appl
31	48.5	21.6	1649	US-09-535-008-75	Sequence 75, Appl
32	48.5	21.6	1650	US-09-535-008-71	Sequence 71, Appl
33	48.5	21.6	1678	US-09-535-008-69	Sequence 69, Appl
34	48.5	21.6	1679	US-09-535-008-65	Sequence 65, Appl
35	48.5	21.6	1681	US-09-535-008-77	Sequence 77, Appl
36	48.5	21.6	1682	US-09-535-008-73	Sequence 73, Appl
37	48	21.3	221	US-09-134-001C-3253	Sequence 3253, Ap
38	48	21.3	525	US-08-356-340-2	Sequence 2, Appl
39	48	21.3	525	US-08-786-555-2	Sequence 2, Appl
40	48	21.3	1105	US-08-999-774A-2	Sequence 2, Appl
41	48	21.3	2512	US-08-801-263A-9	Sequence 9, Appl
42	48	21.3	2512	US-09-102-248-9	Sequence 9, Appl
43	47.5	21.1	405	US-09-413-574-2	Sequence 2, Appl
44	47.5	21.1	614	US-09-832-496-2	Sequence 2, Appl
45	47.5	21.1	614	US-09-832-614A-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-126-420A-18
; Sequence 18, Application US/09126420A
; Patent No. 6376753
;
GENERAL INFORMATION:
; APPLICANT: BATARD, YANNICK
; APPLICANT: ROBINEAU, TIBURCE
; APPLICANT: DURST, FRANCIS
; APPLICANT: WERCK-REICHAUT, DANIELE
; APPLICANT: DIDIERJEAN, LUC
;
TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
TITLE OF INVENTION: TUBEROSOS AND ITS APPLICATIONS AS BIOCATALYST IN
TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
FILE REFERENCE: 03715.0032
CURRENT APPLICATION NUMBER: US/09/126,420A
CURRENT FILING DATE: 1996-07-30
PRIOR APPLICATION NUMBER: 60/054,351
PRIOR FILING DATE: 1997-07-31
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 504
TYPE: PRT
; ORGANISM: Solanum melongena
;
US-09-126-420A-18
Query Match          24.2%  Score 54.5;  DB 4;  Length 504;
Best Local Similarity 30.6%  Pred. No. 16;
Matches 15;  Conservative 3;  Mismatches 10;  Indels 21;  Gaps 2;
                                ||| |
OY 10 PPGPGVAIFGCG-----GPVWPQFTQSPYTLV 39
DB 35 PPGPGPLIFGKMFELGTEPKKNAVLRKQGPVLV--LKEGSTYIMV 81
;
RESULT 2
US-08-944-483-47
; Sequence 47, Application US/08944483
; Patent No. 6232456
;
GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLAAS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
;
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
```

```

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6232456e
US-08-944-483-47

Query Match 23.6%; Score 53; DB 4; Length 251;
Best Local Similarity 36.2%; Pred. No. 12;
Matches 17; Conservative 6; Mismatches 14; Indels 10; Gaps 3;

OY 1 VANFL-LCLPTG-----GQVALFGGPPVWQPTQSMPTLYVVK 41
| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 115 VANFVMPICLPESQOGAMVIVSGW---KQFLQRPETLMEIE 157

RESULT 3
US-09-088-651-2
Sequence 2, Application US/09088651
Patent No. 6165771
GENERAL INFORMATION:
APPLICANT: BURGESS, NICOLA A.
APPLICANT: CLINKENBEARD, HELEN E.
APPLICANT: SOUTHAN, CHRISTOPHER D.
TITLE OF INVENTION: NOVEL COMPOUNDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,651
FILING DATE: JUNE 1, 1998

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9712088.5
FILING DATE: 10-JUNE-1997
APPLICATION NUMBER: EP 97308295.1
FILING DATE: 17-OCT-1997
APPLICATION NUMBER: GB 9803650.2
FILING DATE: 20-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH30358
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-088-651-2

Query Match      23.6%, Score 53; DB 4; Length 304;
Best Local Similarity 45.2%; Pred. No. 14;
Matches 14; Conservative 2; Mismatches 13; Indels 2; Gaps 2

QY    10 PTGGPGVAIFGCGPVPMPQFTQSMPTLVVY 40
      ||||| :||||| | :|
Db     241 PT-GPGCG-WGLGPVWVWLQTEPPSVLIRV 269

RESULT 4
US-08-630-915A-34
Sequence 34, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, NO. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 34:

```


US-08-782-396-2

```

; FILING DATE: 12-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,070C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-070C-10

Query Match 22.2% Score 50; DB 3; Length 281;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 14; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

QY 3 NRFLC---LPTGGPGVAI-FGGGVPWPQFTQSM 33
Db 33 NAVLPSCSYTLPTSGTLVPMCMGKFCPWSQCTNEL 67

RESULT 14
US-08-829-525-10
Sequence 10, Application US/08829525
Patent No. 6084083
GENERAL INFORMATION:
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/829,525
FILING DATE: 28-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,583
FILING DATE: 01-MAR-1996
APPLICATION NUMBER: US 08/487,748
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/398,633
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-081
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-829-525-10

Query Match 22.2% Score 50; DB 3; Length 281;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 14; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

QY 3 NRFLC---LPTGGPGVAI-FGGGVPWPQFTQSM 33
Db 33 NAVLPSCSYTLPTSGTLVPMCMGKFCPWSQCTNEL 67

RESULT 15
US-08-609-583A-10
Sequence 10, Application US/08609583A
Patent No. 6204371
GENERAL INFORMATION:
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,583A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,748
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/398,633
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-609-583A-10

Query Match 22.2% Score 50; DB 4; Length 281;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 14; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

QY 3 NRFLC---LPTGGPGVAI-FGGGVPWPQFTQSM 33
Db 33 NAVLPSCSYTLPTSGTLVPMCMGKFCPWSQCTNEL 67

Search completed: May 20, 2003, 16:44:46
Job time: 20.9231 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:49:49 ; Search time 20.2747 Seconds
(without alignments)
200.549 Million cell updates/sec

Title: US-09-869-155-19

Perfect score: 225
Sequence: 1 VANRFLICPTGPGVAIFG.....GPVMPQFTQSMPTLVVK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US05_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	27.6	291	9	US-09-738-626-5062
2	54	24.0	1220	9	US-09-736-968A-14
3	54	24.0	1220	10	US-09-736-969A-14
4	54	24.0	1220	10	US-09-736-969A-14
5	54	24.0	2090	9	US-09-736-968A-107
6	54	24.0	2090	10	US-09-736-969A-93
7	54	24.0	2090	10	US-09-736-969A-90
8	53.5	23.8	73	10	US-09-867-550-1298
9	53.5	23.8	91	10	US-09-864-761-41973
10	53.5	23.8	684	10	US-09-823-240-9
11	53	23.6	81	9	US-10-023-282-740
12	53	23.6	114	9	US-10-023-282-739
13	53	23.6	136	9	US-09-764-891-4761
14	53	23.6	235	9	US-10-023-282-318
15	53	23.6	235	9	US-10-023-282-738
16	53	23.6	361	9	US-10-149-819-10
17	53	23.6	679	10	US-09-874-198-6
18	53	23.6	679	10	US-09-874-238-6
19	52.5	23.3	441	10	US-09-879-957-34

20	52.5	23.3	519	9	US-10-029-180-8	Sequence 8, Appl
21	52.5	23.3	600	9	US-10-029-180-120	Sequence 120, App
22	52.5	23.3	611	9	US-10-029-180-119	Sequence 119, App
23	52.5	23.3	635	12	US-10-155-059-3	Sequence 3, Appl
24	51	22.7	122	9	US-10-102-806-522	Sequence 522, App
25	50.5	22.4	136	10	US-09-948-018-6	Sequence 6, Appl
26	50.5	22.4	355	10	US-09-948-018-2	Sequence 2, Appl
27	50.5	22.4	380	10	US-09-948-018-36	Sequence 36, Appl
28	50	22.2	110	10	US-09-864-761-40789	Sequence 40789, A
29	50	22.2	218	9	US-10-097-065-230	Sequence 230, App
30	50	22.2	180	12	US-10-155-059-7	Sequence 7, Appl
31	50	22.2	281	9	US-10-004-633-10	Sequence 10, Appl
32	50	22.2	414	9	US-10-125-635A-406	Sequence 406, App
33	50	22.2	414	9	US-09-938-864-406	Sequence 406, App
34	50	22.2	636	10	US-09-732-384-10	Sequence 10, Appl
35	49.5	22.0	245	9	US-10-153-668-92	Sequence 92, Appl
36	49.5	22.0	897	9	US-10-099-895-1	Sequence 1, Appl
37	48.5	21.6	110	10	US-09-732-665-12	Sequence 12, Appl
38	48.5	21.6	152	10	US-09-732-665-11	Sequence 11, Appl
39	48.5	21.6	490	9	US-10-029-180-76	Sequence 76, Appl
40	48.5	21.6	635	9	US-09-877-650-15	Sequence 15, Appl
41	48.5	21.6	635	9	US-10-166-332A-2	Sequence 2, Appl
42	48.5	21.6	625	10	US-09-871-856-15	Sequence 15, Appl
43	48.5	21.6	1647	9	US-09-824-574-4	Sequence 4, Appl
44	48.5	21.6	1724	9	US-09-964-899-43	Sequence 43, Appl
45	48	21.3	75	10	US-09-864-761-37251	Sequence 37251, A

ALIGNMENTS

RESULT 1

US-09-738-626-5062
; Sequence 5062, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5062
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5062

Query Match 27.6% Score 62; DB 9; Length 291;

Best Local Similarity 47.8%; Pred. No. 4.8;

Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 6 LLIPTGPGVAIFGPGVMPQ 28

DB 163 LIIPTGPGVAIFGPGVMPQ 185

```

CURRENT APPLICATION NUMBER: US/09/736,969A
CURRENT FILING DATE: 2000-12-13
PRIOR APPLICATION NUMBER: US 60/160,860
PRIOR FILING DATE: 1999-10-21
PRIOR APPLICATION NUMBER: US 60/162,498
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: US 60/170,453
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 60/176,195
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 60/182,296
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: US 09/547,276
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,267
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,460
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,527
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,528
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 09/687,837
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,503
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,508
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,539
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,543
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 153
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 14
LENGTH: 1220
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human CLASP-3
US-09-736-969A-14

Query Match      24.0%: Score 54; DB 10; Length 1220;
Best Local Similarity 47.8%: Pred. No. 2.2e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY      11 TGGPGCAIFGCGPYWPMQFTQSM 33
      ||| | : | | : ||:|
Db      62 TGGPKAPMGSNPSPASSTQAM 84

RESULT 4
US-09-736-960-14
: Sequence 14, Application US/09736960
: Patent No. US20020102267A1
: GENERAL INFORMATION:
: APPLICANT: Lu, Peter
: APPLICANT: Garman, Jonathan David
: APPLICANT: Candia III, Albert Frederick
: TITLE OF INVENTION: CLASP-5 Transmembrane Protein
: FILE REFERENCE: 020054-000511US
: CURRENT APPLICATION NUMBER: US/09/736,960
: CURRENT FILING DATE: 2001-09-20
: PRIOR APPLICATION NUMBER: US 60/160,860
: PRIOR FILING DATE: 1999-10-21
: PRIOR APPLICATION NUMBER: US 60/162,498
: PRIOR FILING DATE: 1999-10-29
: PRIOR APPLICATION NUMBER: US 60/170,453
: PRIOR FILING DATE: 1999-12-13
: PRIOR APPLICATION NUMBER: US 60/176,195
: PRIOR FILING DATE: 2000-01-14
: PRIOR APPLICATION NUMBER: US 60/182,296
: PRIOR FILING DATE: 2000-01-14

```

```

; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-3
US-09-736-960-14
```

```

Query Match          24.0%; Score 54; DB 10; Length 1220;
Best Local Similarity 47.8%; Pred. No. 2.2e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 11 TGGPGVAIFGCGPVPWPQFTQSM 33
      ||||| | : | | | : ||| |
Db 62 TGGPKAAPMGNSNPSAESSTQAM 84
```

```

RESULT 5
US-09-736-968A-107
; Sequence 107, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-00061105
; CURRENT APPLICATION NUMBER: US/09/736,968A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
```

```

; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 107
; LENGTH: 2090
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-3
US-09-736-968A-107
```

```

Query Match          24.0%; Score 54; DB 9; Length 2090;
Best Local Similarity 47.8%; Pred. No. 3.8e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 11 TGGPGVAIFGCGPVPWPQFTQSM 33
      ||||| | : | | | : ||| |
Db 933 TGGPKAAPMGNSNPSAESSTQAM 955
```

```

RESULT 6
US-09-736-969A-93
; Sequence 93, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-00041105
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 153
```

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 93
LENGTH: 2090
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human CLASP-3
US-09-736-969A-93

Query Match 24.0%; Score 54; DB 10; Length 2090;
Best Local Similarity 47.8%; Pred. No. 3.8e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 11 TGGPGVAIFGCGVPMPQFTQSM 33
||||| : : : : :
Db 933 TGGPKAAPMGSNPSPSAESTQAM 955

RESULT 7
US-09-736-960-90
Sequence 90, Application US/09736960
Patent No. US20020102267A1
GENERAL INFORMATION:
APPLICANT: Lu, Peter
APPLICANT: Garman, Jonathan David
APPLICANT: Candia III, Albert Frederick
APPLICANT: Arbor Vita Corporation
TITLE OF INVENTION: CLASP-5 Transmembrane Protein
FILE REFERENCE: 020054-000511US
CURRENT APPLICATION NUMBER: US/09/736,960
CURRENT FILING DATE: 2001-09-20
PRIOR FILING DATE: 1999-10-21
PRIOR APPLICATION NUMBER: US 60/160,860
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: US 60/170,453
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 60/176,195
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 60/182,296
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: US 09/547,276
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,267
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,460
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,527
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/196,528
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 09/687,837
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,503
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,508
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,539
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 60/240,543
NUMBER OF SEQ ID NOS: 134
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 90
LENGTH: 2090
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human CLASP-3
US-09-736-960-90
Query Match 24.0%; Score 54; DB 10; Length 2090;
Best Local Similarity 47.8%; Pred. No. 3.8e+02;

Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
OY 11 TGGPGVAIFGCGVPMPQFTQSM 33
||||| : : : : :
Db 933 TGGPKAAPMGSNPSPSAESTQAM 955

RESULT 8
US-09-867-550-1298
Sequence 1298, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: US/09/867,550
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1298
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-550-1298

Query Match 23.8%; Score 53.5; DB 10; Length 73;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 13; Conservative 6; Mismatches 11; Indels 9; Gaps 2;

OY 10 PTGPGVAIFGCG-----PVPMPQFTQSMPTLVVV 40
| | | | | : : : : :
Db 17 PAGGSSGSSVSGNGRGKARHPVPVWQFT-ALPVSSCVI 54

RESULT 9
US-09-864-761-41973
Sequence 41973, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 41973
LENGTH: 91
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL135747.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
OTHER INFORMATION: SWISSPROT HIT: P24054, EVALUATE 1.20e+00
OTHER INFORMATION: EST_HUMAN HIT: BF330102.1, EVALUATE 2.00e-22
US-09-864-761-41973

Query Match 23.8%; Score 53.5; DB 10; Length 91;
Best Local Similarity 52.6%; Pred. No. 15;

Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 20 GCGVPMPOFTQSMY-VTL 37
||| |||::|::|::|
Db 16 GCGVCMVPRFSEPDPSHTL 34

RESULT 10
US-09-823-240-9
Sequence 9, Application US/09823240
Patent No. US20020048813A1
GENERAL INFORMATION:
APPLICANT: Frank B. Gertler
APPLICANT: James E. Bear
APPLICANT: Jurgen Wehlend
APPLICANT: Joseph Loureiro
TITLE OF INVENTION: Methods and Products for Regulating Cell
FILE REFERENCE: M0656/7064 (HCL)
CURRENT APPLICATION NUMBER: US/09/823,240
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/194,564
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 684
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-823-240-9

Query Match 23.8%; Score 53.5; DB 10; Length 684;
Best Local Similarity 52.2%; Pred. No. 1.3e+02;

Matches 12; Conservative 1; Mismatches 7; Indels 3; Gaps 1;
QY 13 GCGVAFGCGVPMPOFTQSMY 35
||| ||| ||| ||| |||
Db 351 GPG---YGGPVPPPPQQAENPY 370

RESULT 11
US-10-023-282-740
Sequence 740, Application US/10023282
Publication No. US2003092893A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/10/023,282
PRIOR FILING DATE: 2001-12-20
EARLIER APPLICATION NUMBER: 09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06

```

: EARLIER APPLICATION NUMBER: 60/048,917
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,949
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,974
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,883
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,897
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,898
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,962
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,963
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,877
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,878
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/070,923
: EARLIER FILING DATE: 1997-12-18
: EARLIER APPLICATION NUMBER: 60/092,921
: EARLIER FILING DATE: 1998-07-15
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: NUMBER OF SEQ ID NOS: 1227
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 740
: LENGTH: 81
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-023-282-740

```

```

Query Match      23.6%; Score 53; DB 9; Length 81;
Best Local Similarity 44.4%; Pred. No. 16;
Matches 12; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

```

```

QY      12 GGGVAIFGCGPVPQPT--QSMRYT 36
      111: 11111 11111
Db       8 GPGGNNMGPGGRRPMPNTNNSIPYS 34

```

```

RESULT 12
US-10-023-282-739
: Sequence 739, Application US/10023282
: Publication No. US20030092893A1
: GENERAL INFORMATION:
: APPLICANT: Young et al.
: TITLE OF INVENTION: 207 Human Secreted Proteins
: FILE REFERENCE: P2007P1
: CURRENT APPLICATION NUMBER: US/10/023,282
: CURRENT FILING DATE: 2001-12-20
: EARLIER APPLICATION NUMBER: 09/205,258
: EARLIER FILING DATE: 1998-12-04
: EARLIER APPLICATION NUMBER: PCT/US98/11422
: EARLIER FILING DATE: 1998-06-04
: EARLIER APPLICATION NUMBER: 60/048,885
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,375
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,881
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,880
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,896
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,020
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,876
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,895
: EARLIER FILING DATE: 1997-06-06

```

```

: EARLIER APPLICATION NUMBER: 60/048,884
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,894
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,971
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,964
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,882
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,899
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,892
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,915
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,019
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,970
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,972
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,916
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,373
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,875
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,374
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,917
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,949
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,974
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,883
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,897
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,898
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,962
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,963
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,877
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,878
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/070,923
: EARLIER FILING DATE: 1997-12-18
: EARLIER APPLICATION NUMBER: 60/092,921
: EARLIER FILING DATE: 1998-07-15
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: NUMBER OF SEQ ID NOS: 1227
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 739
: LENGTH: 114
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-023-282-739

```

```

Query Match      23.6%; Score 53; DB 9; Length 114;
Best Local Similarity 44.4%; Pred. No. 22;
Matches 12; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

```

; EARLIER APPLIC

Query Match	23.68;	Score 53;	DB 9;	Length 235;
Best Local Similarity	44.48;	Pred. No. 49;		

```
Matches 12; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
QY 12 GGPVATFGGPPVPOFT--QSMPT 36
| | | | | | | | | | | | | | | |
Db 81 GMPGMMNGGGGRPPNPTNNSIPYS 107

RESULT .5
US-10-023-282-738
; Sequence 738, Application US/10023282
; Publication No. US2003002893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; EARLIER FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 738
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-738

Query Match 23.6%; Score 53; DB 9; Length 235;
Best Local Similarity 44.4%; Pred. No. 49;
Matches 12; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
QY 12 GGPVATFGGPPVPOFT--QSMPT 36
| | | | | | | | | | | | | | | |
Db 81 GMPGMMNGGGGRPPNPTNNSIPYS 107

Search completed: May 20, 2003, 18:00:04
Job time : 21.2747 secs
```